ROBUST DETECTION AND CLASSIFICATION OF REGIONAL SEISMIC SIGNALS USING A TWO MODE/TWO STAGE CASCADED ADAPTIVE ARMA (CAARMA) MODEL

by

Bonnie Schnitta-Israel

Sponsored by
Advanced Research Projects Agency (DOD)
ARPA Order No. 4787

Monitored by AFOSR Under Contract #F49620-83-C-0137

March 1985

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: unclassified	1b. RESTRICTIVE MARKINGS: None
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): TR-159	5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 85-0395
6a. NAME OF PERFORMING ORGANIZATION: KLD Associates, Inc. / 6b. OFFICE SYMBOL (If applicable):	7a. NAME OF MONITORING ORGANIZATION: Air Force Office of Scientific Research
6c. ADDRESS (City, State and ZIP Code): 300 Broadway, Huntington Station, NY 11746	7b. ADDRESS (City, State and ZIP Code): Bolling Air Force Base, DC 20332
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: AFOSR / 8b. OFFICE SYMBOL (If applicable): NP	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: Contract No. F49620-83-C-0137
8c. ADDRESS (City, State and ZIP Code): Building 410, Bolling AFB, DC 20332 use Cover	10. SOURCE OF FUNDING NOS. — PROGRAM ELEMENT NO.: 62714E / PROJECT NO.: (DARPA) 4787 / TASK NO.: 01 / WORK UNIT NO.: N/A
11. TITLE (Include Security Classification): Using a Hybrid Adaptive Processing System, TRAPS I	
12. PERSONAL AUTHOR(S): Bonnie Schnitta-Israel	
13a. TYPE OF REPORT: Final / 13b. TIME COVERED: FROM 83 July 15 TO 84 July 14	14. DATE OF REPORT (Yr., Mo., Day): 1985 March / 15. PAGE COUNT: 167
16. SUPPLEMENTARY NOTATION: cont'd	

17. COSATI CODES		
FIELD	GROUP	SUB. GR.

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

Adaptive Autoregressive (AR), Adaptive Autoregressive Moving Average (ARMA); Linear Predictive Residual; (LPR) Pole Migration; Slayed Adaptive AR Model; Time Delay Estimation Maximum Likelihood.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The blanket objective of the four tasks of the research proposed and completed under this contract was to tailor TRAPS (TRansient Acoustic Processing System) to the seismic event monitoring application thus providing a powerful tool for the nuclear monitoring problem. TRAPS is a three stage process. Stage One is mathematically designed to perform noise elimination on each receiver component. The output of Stage One is then beamformed and input into Stage Two, which performs signal enhancement. The information resulting from Stage One/Two is used in Stage Three of TRAPS to calculate the location and classification of the detected event. Therefore, the achieved goals of this contract's research will be detailed in this report by relating the refinements to these stages, resulting from the contract analysis.

While the original TRAPS, as pictured in Figure 1, succeeded in doubling the range of detection in a previous seismic study (1), it did so with a standard beamformer. Task 1 was

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION: Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL: DR HENRY R. RADOSKI	22b. TELEPHONE NUMBER (Include Area Code): 202/767-4906 / 22c. OFFICE SYMBOL: NP

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

SECURITY CLASSIFICATION OF THIS PAGE

thus to ascertain if an adaptive beamformer (AB) incorporated into TRAPS would produce any gains. Task 1 research efforts determined TRAPS' Stage One followed by a standard beamformer was superior to an AB, yet an AB strategically placed between TRAPS' Stage One and Stage Two did provide some processing gain.

The goal of Task 2 was to investigate and incorporate into TRAPS the relevant background noise statistics for increased TRAPS' efficacy. Task 2 research efforts did result in the formulation of algorithm refinements to TRAPS. One such refinement to TRAPS was the upgrade of the original adaptive autoregressive (AR) model of TRAPS to a unique cascaded adaptive ARMA (CAARMA) model, which provided significant improvement to TRAPS. Anti jamming strategies to prevent deterioration to TRAPS by local interferences, etc., were then included into TRAPS by way of a Slaved Stage One. Additionally, the order and adaptation time of Stage One and Two of TRAPS was varied to improve TRAPS' ability to remove bias in the data.

The Task 3 goal was to decrease the false alarm rate of TRAPS and to enhance FK space definition. It was proposed to accomplish Task 3 goals by means of a signature analysis (SA). Since the coefficients, poles and zeros, etc. generated by the CAARMA process completely characterize the shape of the power spectrum and thus an event, this SA characterization mode was composed of the weighted prediction error residual (WPR) of each event in combination with the event's weighted prediction coefficients (WPC). The SA allowed the achievement of several advancements to TRAPS. First, full waveform or total energy was appropriately classified into its various wave components. This allowed utilization of all available information of an event for improved detection. Once an event was identified, the poles of the CAARMA model were uniquely applied to estimation of time delay among phones. The resulting estimation was input to a feedback maximum likelihood scheme. The culmination of these research efforts produced a powerful processing system, TRAPS II.

Finally, Task 4 examined TRAPS for its possibility of a compact digital processing system. Analysis indicated TRAPS and TRAPS I (See Figures 1 and 2) were capable of real time processing. TRAPS II, as depicted in Figure 3, would require an array processor for real time capability.

FINAL REPORT

ROBUST DETECTION AND CLASSIFICATION OF REGIONAL SEISMIC SIGNALS USING A TWO MODE/TWO STAGE CASCADED ADAPTIVE ARMA (CAARMA) MODEL

ARPA Order No:	4787
Program Code:	3A10
Name of Contractor:	KLD Associates, Inc.
Effective Data of Contract:	15 July 1983
Contract Expiration Date:	14 July 1984
Amount of Contract Dollars:	$99,887
Principal Investigator:	Bonnie Schnitta-Israel (516) 324-1300
Program Manager:	Daniel E. Magnus (516) 549-9803
Contract No.:	F4960-83-C-0137
Title of Work:	Robust Detection and Classification of Regional Seismic Signals

Sponsored by:

Advanced Research Projects Agency (DOD)
ARPA Order No. 4787
Monitored by AFOSR under Contract No. F49620-83-C-0137

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont'd)

Abstract

The blanket objective of the four tasks of the research proposed and completed under this contract was to tailor TRAPS (TRansient Acoustic Processing System) to the seismic event monitoring application thus providing a powerful tool for the nuclear monitoring problem. TRAPS is a three stage process. Stage One is mathematically designed to perform noise elimination on each receiver component. The output of Stage One is then beamformed and input into Stage Two, which performs signal enhancement. The information resulting from Stage One/Two is used in Stage Three of TRAPS to calculate the location and classification of the detected event. Therefore, the achieved goals of this contract's research will be detailed in this report by relating the refinements to these stages, resulting from the contract analysis.

While the original TRAPS, as pictured in Figure 1, succeeded in doubling the range of detection in a previous seismic study (1), it did so with a standard beamformer. Task 1 was thus to ascertain if an adaptive beamformer (AB) incorporated into TRAPS would produce any gains. Task 1 research efforts determined TRAPS' Stage One followed by a standard beamformer was superior to an AB, yet an AB strategically placed between TRAPS' Stage One and Stage Two did provide some processing gain.

The goal of Task 2 was to investigate and incorporate into TRAPS the relevant background noise statistics for increased TRAPS' efficacy. Task 2 research efforts did result in the formulation of algorithm refinements to TRAPS. One such refinement to TRAPS was the upgrade of the original adaptive autoregressive (AR) model of TRAPS to a unique cascaded adaptive ARMA (CAARMA) model, which provided significant improvement to

TRAPS. Anti jamming strategies to prevent deterioration to TRAPS by local interference, etc., were then included into TRAPS by way of a Slaved Stage One. Additionally, the order and adaptation time of Stage One and Two of TRAPS was varied to improve TRAPS' ability to remove bias in the data.

The Task 3 goal was to decrease the false alarm rate of TRAPS and to enhance FK space definition. It was proposed to accomplish Task 3 goals by means of a signature analysis (SA). Since the coefficients, poles and zeros, etc. generated by the CAARMA process completely characterize the shape of the power spectrum and thus an event, this SA characterization mode was composed of the weighted prediction error residual (WPR) of each event in combination with the event's weighted prediction coefficients (WPC). The SA allowed the achievement of several advancements to TRAPS. First, full waveform or total energy was appropriately classified into its various wave components. This allowed utilization of all available information of an event for improved detection. Once an event was identified, the poles of the CAARMA model were uniquely applied to estimation of time delay among phones. The resulting estimation was input to a feedback maximum likelihood scheme. The culmination of these research efforts produced a powerful processing system, TRAPS II.

Finally, Task 4 examined TRAPS for its possibility of a compact digital processing system. Analysis indicated TRAPS and TRAPS I (see Figures 1 and 2) were capable of real time processing. TRAPS II, as depicted in Figure 3, would require an array processor for real time capability.

1.0 SUMMARY

TRAPS (TRansient Acoustic Processing System) is a three stage process. Stage One is a hybrid adaptive filter, which is mathematically designed to perform noise elimination on each phone. The output of Stage One is then beamformed and input into Stage Two, another hybrid adaptive filter which serves to enhance the signal. The information resulting from Stage One/ Two is used in Stage Three of TRAPS to calculate the range, bearing and classification of the detected event. The blanket objective of the research proposed and completed under this contract was to improve the existing detection and localization capability of TRAPS by means of tailoring each of TRAPS' three Stages to the seismic event monitoring application. This broad objective was broken down into 4 specific Tasks.

While the original TRAPS succeeded in doubling the range of detection in a previous seismic study (1), it did so with a standard beamformer, as opposed to an adaptive beamformer (AB). Task 1 was thus to ascertain the most effective mode of AB incorporation into TRAPS and specify any gains resulting from the incorporation. Task 1 research efforts determined TRAPS' Stage One followed by a standard beamformer was superior to an AB, yet an AB strategically placed between TRAPS' Stage One and Stage Two did provide some processing gain. Section 2 compares the effectiveness of the original TRAPS to TRAPS with an AB and to other spectral estimators as illustrated by means of a

normalized frequency procedure.

The goal of Task 2 was to analyze the environment within which TRAPS was to function. The knowledge of the relevant background noise statistics accumulated by this study would then be incorporated into TRAPS for its improvement. Task 2 efforts did inspire the formulation of two algorithm refinements to TRAPS. The first refinement addressed the signal cancellation phenomena which is a consequence of adaptive interaction between the desired signal and interference, certain possible false signals generated that defeat or partially deteriorate the adaptive filter, etc. As a result, a sophisticated slaved algorithm was built into Stage One of TRAPS as an anti-jammer to prevent local interference, etc. The second change to TRAPS was influenced by the knowledge that the autoregressive moving average (ARMA) model often provides better spectral estimates than the more specialized AR models. Thus the original adaptive AR model of TRAPS was transformed to an adaptive ARMA model, which could provide source parameterization improvement in a low SNR environment, while maintaining real time processing criteria. These refinements are presented in Section 3.2. The resulting model was then upgraded by a unique cascade approach to the problem. This cascaded adaptive ARMA (CAARMA) model, which is detailed in Section 3.3 and 3.4 provided excellent results. Additionally, the order and adaptation time of Stage One and Two of TRAPS was varied to improve its ability to remove various bias in the data, as discussed in

Section 3.5. The end result was a processing system that had excellent results, as presented in Section 3.6.

The goal of Task 3 was to decrease the false alarm rate of TRAPS and to enhance FK space definition by means of a signal characterization mode. In order to demonstrate the logic for the direction taken in Task 3 it should be noted that the coefficients, poles and zeros, etc., generated by the CAARMA process completely characterize the shape of the power spectrum and thus an event. Thus the similarity of one event to another can be verified by comparing the parameters which result from Stage One and Two of TRAPS. The signature of one event was then compared to the signature of another event by means of the weighted prediction error residual (WPR) of each event. These WPR's in combination with the weighted prediction coefficients (WPC) served as the foundation for a statistical/structural classification system, or pattern recognition algorithm (PRA), as presented in Section 4.3.

The PRA allowed the achievement of several improvements to TRAPS. First, such a PRA is able to discriminate between various types of explosions and earthquakes. Also, the information resulting from the PRA isolated and then classified the various waves of an event. This PRA was also capable of identifying phase irregularities that characterized specific field data. This technique uses the full waveform or total energy, thus allowing utilization of all available information of an event for determination of detection, location and classification.

There is no penalty built into the system for "missing" information. The advantage lies in the fact that whatever information is present is appropriately weighted to provide maximum efficacy. Once an event was identified the poles of the CAARMA model were uniquely used for time delay estimation. This innovative approach was initiated by earlier results that the nature of the pole migration served as a robust event detector. This time delay estimation was input to a feedback maximum likelihood scheme. The culmination of these research efforts, which produced a powerful processing system, is discussed in Section 4.4.

Finally, Task 4 examined TRAPS for incorporation into a compact digital system. Once the reference sets of the PRA were established, it was found that human intervention was not necessary. The only limit to the TRAPS resulting from the above research was that real time processing could occur only with the aid of an array processor.

2.0 TASK 1: EVALUATION OF TRANSIENT ACOUSTIC PROCESSING SYSTEM (TRAPS) FOR PROCESSING GAINS DUE TO AN ADDED ADAPTIVE BEAMFORMER

2.1 Introduction to the Fundamental Hybrid Model of TRAPS

Section 2.1 shall present the mathematical foundation of the original TRAPS and its relationship to the physics of the nuclear monitoring problem. This process will enable the reader to then comprehend the reason for TRAPS' initial success in seismic event detection and location (1), the logic behind the various improvements to TRAPS and finally insight into the unique capability and consequential success of the improved TRAPS, TRAPS I for the nuclear monitoring application. It is important to understand that TRAPS is in essence a three stage algorithm. These stages are (as illustrated in Figure 1):

- Stage One - Noise Elimination
- Stage Two - Signal Enhancement
- Stage Three - Range and Bearing Estimation

This report shall first concentrate on some of the terminology and theory of the adaptive autoregressive (AR) model which Stage One and Two of the original TRAPS was based. Let $x(n) = x(t_n) = x(n\Delta t)$; $n = 0, 1, 2, \ldots$, denote a particular seismic wave component at the geophone location. $x(n)$ is then

considered to have originated from a source located some unknown distance from the geophones. The received sequence is assumed to satisfy a difference equation of the sort:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots + \phi_p Z_{t-p} + \varepsilon_t \qquad 2.1$$

where the ϕ_p are constants, and the forcing function ε_t is a stochastic variable with Gaussian statistics during an interval $T_o \leq n\Delta t \leq T_1$ and is zero for $n\Delta t$ outside this interval. Given that $(T_o, T_1) = (-\infty, \infty)$ and that ε_t is stationary this model would be called an autoregressive process of order p (2).

For this adaptive AR model of TRAPS the finite duration stochastic function, ε_t, can be considered to represent the unstructured energy in the wave, and the coefficients, ϕ_p, can be thought of as representing the filtering effect of the earth and the initial source structure. The received wave is modeled as the output of a linear recursive system which at some time T_o is excited by Gaussian noise, and at some time T_1 is de-excited. While this is still an autoregressive process of order p (3), the sequences of the same type as eqn (2.1) in hybrid adaptive form is referred to as "transient autoregressive." The order to the transient autoregressive process is critical and various methods of selection have been studied extensively (4-7). The method chosen for identifying the order of the process and the order we chose for nuclear monitoring is addressed in Section 3.5. This allows us to consider p as known for this

discussion. However, we do not know either $[\phi_p]$ or $[\varepsilon_t]$ and in order to circumvent this difficulty we employ a hybrid adaptive filtration technique. We summarize here the adaptive least mean square (LMS) filter originally used in these data studies. The theoretical basis of the adaptive LMS algorithm is an extension of work by Widrow (8) and Griffiths (9), which can be referenced for further detail.

The adaptive filter constructs the sequence of prediction filter coefficients $[\phi_p(n)]$ which converge to true coefficients as $n \to \infty$, by first predicting a input $\hat{Z}(n)$ at time n from a linear combination of previous values of the actual input $Z(n)$. Thus,

$$\hat{Z}(n) = \sum_{p=1}^{P} \phi_p(n) Z(n-p) \qquad 2.2$$

This allows an error sequence $\varepsilon(n)$,

$$\varepsilon(n) = Z(n) - \hat{Z}(n) \qquad 2.3$$

to be calculated. The set of filter coefficients $\phi_1^*, \phi_z^*, \ldots, \phi_p^*$ which yields the minimum mean-square-error $E[\varepsilon^2(n)]$ is

$$E[\varepsilon^2(n)] = E[[Z(n)-\hat{Z}(n)]^2]. \qquad 2.4$$

This equation is a quadratic function of the filter coefficients and has a unique minimum value which is achieved by

$[\phi_p^*](10)$. The coefficients are then given by the LMS algorithm

$$\phi_p(n+1) = \phi_p(n) + \mu[Z(n)-\hat{Z}(n)]\ Z(n-p) \qquad 2.5$$

where μ is the proportionality constant. The choice of μ determines both the rate of convergence of the processor and the steadystate mean-square-error. Convergence of the algorithm is assured for stationary autoregressive data [Z(n)] of order P provided

$$\mu = \frac{\beta}{P\sigma_Z^2} \qquad 2.6$$

where σ_Z^2 is the input power level and $0<\beta<2$. The adaptation time constant of the algorithm expressed in data points is

$$\tau = \frac{-1}{\mathrm{Ln}(1-\mu\sigma_Z^2)} \cdot \qquad 2.7$$

The method of selecting the adaptation time as well as its effect on the processing gain will be presented in Sections 2.2 and 3.5.

The First Stage of TRAPS uses this adaptive transient AR model to filter out the ambient noise. The ambient noise field is considered to be the sum of a number of travelling waves. The sampled time trace at a fixed spatial point of each of these waves is assumed to be a transient

autoregressive sequence as defined earlier. Therefore, the measured time trace of the ambient noise field at a fixed point is a sequence which is the sum of a number of transient autoregressive sequences. Such a sequence is not in general autoregressive. The following assumptions are made regarding the ambient noise field:

1) there is a dominant (energy wise) AR component
2) this component is stationary or at worst slowly varying
3) an adaptive least mean square (LMS) filter with appropriate parameters converges to (and tracks) this component.

These assumptions are justified empirically by a large number of ambient spectra that we derived from the NORESS data. The assumptions are, moreover, justified physically by virtue of the bandpass property, with respect to wavenumber, of a layered half-space. The bandpass characteristic of the earth channel leads to dominant energy frequencies as well as a temporal extension of signals. The superposition of a number of such signals results in a stationary signal as in the classical analysis of filtered shot noise (11).

The original TRAPS algorithm presented in this section was designed to model all distinct signals which are measured by the geophone array. All wave activity travelling the earth originates from some source, and therefore in this sense all measurements are due to sources of the type we are

interested in. However, as each signal's distance from its origin increases differential absorption, normal spreading and viscous effects cause signals to overlap one another more and are generally of lower individual amplitude and jointly of less pronounced structure. When these signals are received, the net intensity may be significant. Stage One of TRAPS attempts to eliminate as much of this quasi-stationary background as possible by means of a particular adaptive filter.

Let us suppose that we have knowledge of the joint ambient autoregressive equation postulated in item 1 above. We remark that this equation can be obtained adaptively as proposed in item 3 above or by means of an asymptotic surface wave analysis based upon a detailed knowledge of the geographical layering at the test site (12). The measurement may then be written as

$$x(n) = z(n) + u(n) \qquad 2.8$$

where u(n) is the dominant ambient component so that u(n) satisfies the "ambient equation":

$$L\left[\underset{\sim}{u}(n)\right] = w(n) \qquad 2.9$$

and z(n) is the residual component not satisfying this equation (13). The notation $L\left[\underset{\sim}{u}(n)\right]$ denotes the autoregressive operator L acting on the vector of data values $\underset{\sim}{u}(n)$ known at time n.

The variable w is the excitation variable, which by assumption 2 above is essentially white noise.

The noise cancellation proceeds as follows:

Define

$$\varepsilon(n) = L\left[\underset{\sim}{x}(n)\right] \qquad 2.10$$

Then we see

$$\varepsilon(n) = y(n) + w(n) \qquad 2.11$$

where $y(n) \equiv L\left[Z(n)\right]$ is transient autoregressive. That $y(n)$ is transient autoregressive follows from the commutativity of constant coefficient difference operators. The earlier definition of $[T_o, T_1]$ shall be extended very slightly to allow the interval $[T_o, T_1]$ to be a union of such intervals. The definition itself does not require the noise to be homostochastic.

For these LMS filters acting on the individual phones the adaptation parameters must be chosen such that

$$T_A \ll \tau \Delta t \gg T_s \qquad 2.12$$

where T_A is the ambient time scale and T_s is the signal duration. This time scale allows TRAPS to track and then eliminate the ambient noise structure without disturbing the transient signal structure. In summary, Stage One gives us a

remarkable (and computationally fast) SNR improvement, due to the following virtues:

1) Independent adaptive filters implemented for each geophone, K, (and channel) in an array will converge to the ambient operator L_k which is local to that geophone. The resulting noise sequences $w_k(n)$ will be independent.
2) The sequences $\varepsilon_k(n)$ can then be combined in a manner chosen to accentuate the spatially coherent structure of any incoming signals which are rapidly varying with respect to the ambient time scale. The independence of the sequences $w_k(n)$ will produce maximum signal to noise ratio.

Once the first Stage of TRAPS has performed the necessary noise elimination, the filtered data is then processed by Stage Two of TRAPS. Stage Two gives mathemetical priority to signal enhancement. In order to describe an incoming wave of interest we let $y_p(n)$ denote the observed geophone measurement on geophone p. Then

$$z(n) \equiv \sum_{p=1}^{P} y_p(n) \cos(k_1 x_p + k_2 y_p) \qquad 2.13$$

where (x_p, y_p) are the coordinates of the p^{th} geophone and

$$k_1 = \frac{2\pi}{\lambda} \cos \theta \qquad 2.14$$

$$k_2 = \frac{2\pi}{\lambda} \sin \theta \qquad 2.15$$

where λ and θ are wavelength and bearing of an incoming plane wave, respectively. The instantaneous power spectrum of the detected wave,

$H_z(f;n)$ of $z(n)$ at time $t_n = n\Delta t$ and frequency f, is given by

$$H_z(f;n) = \frac{2\Delta t \sigma_z^2}{[h(f;n)]^2} \qquad 2.16$$

$$h(f;n) = 1 - \sum_{m=1}^{M} a_m(n) \exp(-i2\pi fm\Delta t) \qquad 2.17$$

where $a_m(n)$ are adaptive coefficients modeling $z(n)$ as

$$\hat{Z}(n) = \sum_{m=1}^{M} a_m(n)\, Z(n-m) \qquad 2.18$$

$$a_m(n+1) = a_m(n) + \eta z(n-m)[z(n) - \hat{z}(n)] \qquad 2.19$$

For the filters acting on this beamformed data the adaptation time, η, varied from one event to another in order to incorporate the local variance of each event.

The presence of a signal was determined by sharp changes in the value of the spectrum (2.16) within a frequency range acceptable to physical laws that govern the nuclear monitoring situation. Filtration in this narrow band

successfully detected the provided events. The range of detection for previous seismic applications (1) was more than doubled.

Figures 4 and 5 are provided to illustrate just how effective is the first stage and second stage of TRAPS for the NORESS application. Figure 4 is the output of just Stage Two of TRAPS. Figure 5 is the output achieved by using Stage One and Stage Two of TRAPS.

2.2 Mathematical Detail of Task 1 Evaluation Procedure

Task 1 of the Contract was to expand TRAPS to include an adaptive beamformer (AB). It was necessary to evaluate the positioning of the adaptive beamformer into TRAPS so as to determine the position which provided greatest processing gain. The performance criterion used to judge the processing gain of the various positions an AB in TRAPS was a two-sinusoid frequency resolution (14). This method may be best understood by considering two sinusoids with frequencies, f_1 and f_2. The center frequency of these two frequencies, f_c, is $(f_1 + f_2)/2$. The condition for two-sinusoid resolution in this evaluation is defined as the frequency separation, $\Delta f = [f_1 - f_2]$ at which the power spectral density evaluated at the center frequency, $S(f_c)$, is equal to the average of the power spectral densities evaluated at the two sinusoid frequencies, i.e.,

$$S(f_c) = \frac{1}{2}\left[S(f_1) + S(f_2)\right] \qquad 2.20$$

This condition is demonstrated in Figure 6.c. The motivation for the definition of the condition for resolution was to provide a common method that could be applied to the various spectral estimation under investigation. The definition is independent of any property of a specific spectral estimation method. It is also computationally simple to evaluate. Eq.

2.20 does not require exact knowledge of the location of the spectral peak. Also, Eq. 2.20 was beneficial in that the just-resolved condition of Figure 6.c. occurs at a point near the resolution limit, so that a slight shift of f_1 and f_2 toward f_c would merge the peaks into one.

The measure of frequency resolution is thus provided by a dimensionless quantity R, or normalized resolution,

$$R = 2\pi \Delta t \Delta f \qquad 2.21$$

where Δt is the sampling interval in seconds and Δf is the frequency separation in Hz point resolved according to Eq. 2.20. This was viewed to be an optimum definition, since the only frequency variable required in the computer simulations for all spectral estimation techniques was the radian frequency λ, where

$$\lambda = 2\pi f \Delta t \qquad -\pi \leq \lambda \leq \pi \qquad 2.22$$

and f is the true frequency in Hz. Thus, it was unnecessary to specify an actual sample interval Δt or frequency f in order to run the simulations. The condition of resolution occurs when

$$S\left(\frac{\lambda_1 + \lambda_2}{2}\right) = \frac{1}{2}\left[S(\lambda_1) + S(\lambda_2)\right] \qquad 2.23$$

The true resolution in Hz is obtained from the normalized resolution by

$$\Delta f = R/2\pi\Delta t \qquad 2.24$$

For the following comparisons use was made of a modified normalized resolution

$$R_N = 2\pi N\Delta t\Delta f \qquad 2.25$$

plotted as a function of N, the number of data points, or auto-correlation lags. Given the sample interval, Δt and N, plus the value of R_N determined for the spectral estimation technique being evaluated, the frequency resolution in Hz will be given by

$$\Delta f = R_N/2\pi N\Delta t \qquad 2.26$$

Note that smaller values of R_N imply better frequency resolution.

The effect of noise on the resolution of the various spectral methods evaluated was measured by running an ensemble of computer simulations that added white noise to the two sinusoids. During each resolution measurement, the noise was held constant while the frequency separation of the sinusoids was gradually decreased until the resolution criterion was met. For each series of measurements, this procedure was repeated

over independent noise epochs. The initial phases of both sinusoids were fixed at 0°. The center frequency between the sinusoids was at half the foldover frequency, or the center of the Nyquist bandwidth.

The goal of Task 1, as mentioned earlier, was to evaluate whether or not there was a differential gain of TRAPS once expanded to include an adaptive beamformer (AB). Evaluation needed to also demonstrate the various positions of an AB within TRAPS (front end, between Stage One and Two of TRAPS, etc.) and their respective differential gains. Adaptive beamforming in the time domain can basically be categorized into two mathematical concepts, the constrained minimum power and the unconstrained minimum "error." Due to this fact two versions of TRAPS were prepared for AB implantation. The first algorithm was based on Wen-Wu Shen's "A constrained minimum power adaptive beamformer with time-varying adaption rate" (15), which we shall hereafter refer to as Shen's algorithm. Shen's algorithm was chosen due to its ease of implementation into TRAPS and its reported success with seismic data. The second algorithm, which fit into the unconstrained minimum error category, was based on an algorithm recently developed by Bernard Widrow and Kenneth Duvall at Stanford University (16). This algorithm, hereafter referred to as Widrow's algorithm, was implemented into TRAPS for comparison also due to its reported success in seismic applications. The reader is referred to Shen's (15) and Widrow's (16, 17) articles for mathematical

detail of each AB, since the intent of this report is not to duplicate well written previously published algorithms.

To make the Task 1 evaluation, it was believed that a standard power spectral density method should be included in the analysis to serve as a benchmark. A form of the periodogram was chosen since it is a standardized method of power spectral density (18). The modern form of the periodogram is computed as the modulus of a modified Fourier transform.

$$\hat{S}_{per}(f) = \frac{1}{N\Delta t} \left[\Delta t \sum_{n=0}^{N-1} X_n \exp(-j2\pi fn\Delta t)\right]^2 \qquad 2.27$$

for

$$0 \leq f \leq f_{Nyquist}$$

where

$f_{Nyquist}$ = $1/2\Delta t$, the folding, or Nyquist frequency

Δt = sampling interval

N = number of data samples

$N\Delta t$ = total observation time

x_n = $x(n\Delta t)$, the data samples for n=0,..., N-1.

The function described by Eq. 2.27 is a cyclic function of f, repeating its values at intervals of $1/\Delta t$ Hz.

By evaluating Eq. 2.27 at only a set of frequencies spaced equally apart by $1/N\Delta t$, the DFT method of spectrum

analysis is obtained:

$$F_m = \Delta t \sum_{n=0}^{N-1} x_n \exp(-j2\pi nm/N) \qquad 2.28$$

$$\hat{S}_{DFT}(m) = \frac{1}{N\Delta t} |F_m|^2, \quad m = 0, \ldots, N-1 \qquad 2.29$$

The exponential argument is not a function of the sample interval, Δt, due to the particular selection of the frequency spacing. That is, the same number of transform "samples" as data samples are computed. If N is chosen to be 2^r, for r some positive integer, then a computationally efficient algorithm, known as the FFT, may be used to solve for the F_m of Eq. 2.28.

2.3 Technical Results

Figure 7 is the plot of the normalized resolution R_M versus M for mean resolution of a standard FFT (periodogram) and the various TRAPS algorithms resolution at three different SNRs. The SNR is defined as a function of the sinusoid power P_S and the noise variance, $P_N = \sigma_n^2$,

$$\text{SNR(in dB)} = 10 \log P_S/P_N \qquad 2.30$$

A fact to be noted from Figure 7 is that the mean resolution of the periodogram is independent of the SNR. However, the variance about the mean increases with decreasing SNR. Table 1 demonstrates this phenomena for N = 256. The SNR independence is not surprising since the windowing is the key contributing factor that limits the periodogram PSD estimate resolution.

The AR PSD resolution is also a function of the SNR. It can be seen from Figure 7 that the AR technique produces significantly higher resolution than the conventional technique. For example, in Figure 7 for N = 33 and 20db SNR, there is a significant improvement in resolution using the AR technique over the averaged periodogram. This is in agreement with the general findings of other researchers (19, 20).

3.0 TASK 2: EXTENSIONS TO TRAPS FOR NOISE ELIMINATION

3.1 Introduction

The goal of Task 2 was the investigation and incorporation into TRAPS of background noise statistics relevant to the nuclear monitoring problem. Since, as cited in Section 2.1, TRAPS Stage One cancelled all signals which did not belong in the certain class of transients that comprise the seismic events of interest, initial research efforts concentrated on means of improving TRAPS' sophisticated noise elimination stage. The first such improvement was the replacement of the adaptive AR model of the original TRAPS with an adaptive autoregressive moving average (AARMA) model. This allowed the algorithm to better parameterize both the signal and the noise, thus increasing TRAPS' efficacy in a low signal to noise (SNR) environment. The second improvement to TRAPS' Stage One was the expansion of the first Stage to include a reference waveform that slaved the AARMA model. This process not only prevented jamming, but also inhibited the passing to TRAPS' Stage Two noise of similar characteristics to the signals of concern. Both these improvements to TRAPS were detailed in the semi-annual report and shall be briefed in Section 3.2 for the sake of continuity in this final report.

The research efforts in the first six months produced an accurate AARMA model, which was constrained to remain stable.

The accuracy and consequential noise elimination provided by the AARMA model for the majority of the cases is impressive. Extensive analysis in the final six months of the contract showed a weakness in the model, when the noise had a sharp transition in its characteristics. While this weakness was prevalent for only a short time span, it did deteriorate the model. This weakness was eliminated by putting the AARMA into a two mode process. Mode One was as described in Section 3.2. Whenever the error function reflected specific conditions the model switched to Mode Two which was the AARMA mdoel expanded into a cascade structure called CAARMA. Sections 3.3 and 3.4 present this model as well as comparison studies.

Finally, the CAARMA model was tailored to the nuclear monitoring problem. The noise statistics as perceived by an array were evaluated for the relationship to the parameters of a CAARMA model. This research is presented in Section 3.5.

3.2 An Adaptive Autoregressive Moving Average (AARMA) for Improved Detection in Low SNR with Anti-Jamming Strategies

The original TRAPS which was presented in Section 2 was based on an adaptive AR process. This was because a priori knowledge about the generating mechanism of the time series of the original application was not known. However, for the situation of explosion/earthquake monitoring, the generating mechanism, as well as the background noise statistics are available. Cognizance of wave propagation models along with initial evaluation of the NORESS data, indicated that a processing gain would be achieved by upgrading TRAPS to an adaptive ARMA model. This was due to two factors. The first factor is that the AARMA generally improves spectral resolution at low SNR, since the ARMA model often provides better spectral estimates than the more specialized AR or MA models (21). The second factor to be considered in applying the ARMA model to seismic signal processing for improved detection and localization is that as the distance from source to receiver increases, there is the increase in multipath, which may be represented by an ARMA system (22). The ability of the AARMA to improve spectral resolution and to accurately model the NORESS data is innately related to the physics of the problem, which we shall briefly outline at this time.

The unprocessed data from a physical point of view is the response of the system consisting of the geophone array and the media through which the impulsive source is propagated. Thus the received data may be defined as the sum of time-varying responses to sharp impulses, each sharp impulse associated with a media interface that reflects the energy to some degree. The time-varying responses are the wavelets. The interpretation of the received waveform consists of breaking down this elaborate wavelet into its component wavelets. The mathematical solution to the interpretation problem can be understood by first assuming that a given section of raw data trace is additively composed of wavelets. This wavelet shape can be represented by a stable, casual, minimum-delay time function. Thus, each wavelet is assumed to be a one-sided transient that damps with a certain degree of rapidity. Also, knowledge of the arrival time of one wavelet does not allow prediction of the arrival time of another wavelet. It is further assumed that knowledge of the strength of one wavelet cannot predict the strength of another wavelet.

These seismic trace assumptions can now be translated into mathematical notation for discrete time t. First, the fundamental constant wavelet can be represented at discrete, equally spaced, integer-valued times t. Relative to the nature of this wavelet, its shape is given by a set of coefficients, b_t, that are weighted by their respective strengths, i_t. That

is, at time t, the wavelet that arrives then gives the contribution $i_t b_o$; the wavelet that arrived at t-1 gives the contribution $i_{t-1} b_1$; etc. The seismic trace, x_t, at time t, is then the summation of all these contributions. This wavelet complex may then be written as the finite sum

$$x_t = b_o i_t + b_1 i_{t-1} + b_2 i_{t-2} + \ldots = \sum_{s=0}^{\infty} b_s i_{t-s} \qquad 3.1$$

for the arrival times between t_1 and t_2 called the time interval or time gate, which comprises our basic section of seismic trace. This equation includes tails of wavelets with shape, b_t, these wavelets being due to sharp, yet not infinitely sharp, impulses $i_{t_1}-1$, $i_{t_1}-2$,... which occur before time t_1. Without loss of generality, the sharp impulses i_t may be centered so that their mean $E\{i_t\}$ is equal to zero. In the above equation coefficients b_t represent the "dynamics" of the time series, whereas the finite duration stochastic function it represents is the "random" nature of the time series.

It is conceptually easy to imagine this decomposition of a seismogram. It is quite another matter to solve the problem of separating the "dynamic" and the "random" components of a finite section of data. As is well known, state-variable methods and models are efficient means of handling time-varying, non-linear and/or nonstationary effects, multichannel data, and uncertainty (23). A state-variable model is that set of

equations which describe the unique relations between the input, output and state. The state variables in the geophone and layered earth system has a well-defined physical meaning, as just described. It is thus possible to identify the system by the structure of the transfer function. The ARMA model is such a transfer function. An ARMA model can be written in the following notation:

$$x(n) - \sum_{i=1}^{p} a_i x(n-i) = \sum_{i=0}^{q} b_i W(n-i) \qquad 3.2$$

in which x(n) is the modeled time series and W(n) is a related "white" noise series. Taking the z transform ($Z=\exp(i2\pi\Delta tf)$, we see that we have

$$x(Z) = \frac{B(Z)}{A(Z)} W(Z) \qquad 3.3$$

where A(z) and B(z) are p,q degree polynomials in z, and where W(z) is the transform of the series W(n). We consider the observed series x(n) to be the output of a filter with transfer function $H(z) = B(z)/A(z)$ and input W(n).

So as to better understand the system identification by the structure of the transfer function, the transport media may be viewed as made up of layered strata and the seismic wave propagation in this layered medium as reasonably modeled by plane waves. The matching of wave equation jump conditions at each

interface leads to the definition of a reflection coefficient, as discussed earlier in this section. Thus, the response of a system of varied density to a unit spike is given by the equation 3.2, where the a's are related to the reflection coefficients and the b's are functions of the media associated with factors that could be considered as interfering with an accurate event detection. We will consequently categorize the variations in the signal, which contribute to these terms, as noise. The ARMA equations need now to be specified over a finite time and devise a method for solving the unknowns in an environment which corrects for gradual variation of the transient signal. Referring to our earlier discussion of the advantages provided by an adaptive process in Section 2, we propose to use an adaptive ARMA model. If we were concerned only with an adaptive AR process, the equations to be solved would be quadratic in nature and the unknowns could easily be resolved, as discussed in Section 2.1. The adaptive ARMA model is nonlinear and thus requires a method which remains stable within the process of solving the nonlinear equation without requiring excessive computer time. In order to accomplish this, the following hybrid adaptive ARMA model was developed. First, the model in equation 3.2 is rewritten in two stages as follows:

$$\text{I:} \quad e(n) = x(n) - \sum_{i=1}^{P} a_i x(n-i) \qquad 3.4$$

and

$$\text{II:} \quad \sum_{i=0}^{p} b_i W(n-i) = e(n) \qquad 3.5$$

It is possible to initially calculate an estimate, $\hat{e}(n)$, for the series e(n) (2). This leaves the problem of estimating b_i, W_n given equation 3.5. We now estimate b_i and W(n) as those values $\hat{b}_i$, $\hat{W}(n)$ which minimize the total squared error:

$$G_n(\underset{\sim}{b},\underset{\sim}{W}) = \sum [F_n(\underset{\sim}{b},\underset{\sim}{W}) - \hat{e}(n)]^2 \qquad 3.6$$

Subject to the constraints

$$R_p(n) = \sum_{i=0}^{m} W(n-i)W(n-p-i) = \begin{cases} C & p=0 \\ 0 & 1 \leq p \leq s \leq m \end{cases} \qquad 3.7$$

where C is an arbitrary constraint, s is the number of unknown $\underset{\sim}{W}$ and

$$F_n(b,W) = \sum_{i=0}^{q} b_i W(n-i) \qquad 3.8$$

Note that:

1) At this stage $\hat{e}(n)$ is considered given.
2) Any algorithm obtaining $\hat{e}(n)$ must eventually be justified

3) More precision on limits to sums will be made subsequently.

We now recast these equations into a form suitable for numerical treatment.

Suppose that at time $t_n = n\Delta t$ we have observed (estimate) e(n). (We drop the carets for notational convenience.) We compute F from a backward sum

$$F_{n-k} = \sum_{i=0}^{q} b(i+1)W(n-k-i) \qquad 3.9$$

where $k=0,\ldots,N \quad (n>N)$

and $b(i+1) = b_i$, where

$$W(n-k-i) = W_{n-k-i} \quad (n>k+i) \qquad 3.10$$

are FORTRAN arrays representing the unknowns. We represent these unknowns jointly as $\underset{\sim}{\xi}$ (at epoch n):

$$\xi_m = \begin{cases} W(n-m+1) & m \leq s \\ b\ (m-s) & m > s \end{cases} \qquad 3.11$$

where $1 \leq m \leq s + q + 1$ and q+1 is the number of unknown $\underset{\sim}{b}$. We write our problem (at epoch n) as

$$\delta \left\{ \sum_{k=0}^{N} \left(F_{n-k}(\xi) - e_{n-k}\right)^2 + \sum_{p=0}^{s-1} \lambda_p (R_p - C\delta_{po}) \right\} = 0 \qquad 3.12$$

where λ_p are Lagrange multipliers.

The normal equations are:

$$2\sum_{k=0}^{N}\left[(F_{n-k}-e_{n-k})\ \frac{\partial F_{n-k}}{\partial \xi_m}+\sum_{p=0}^{s-1}\lambda_p\ \frac{\partial Rp}{\partial \xi_m}\right]=0 \qquad 3.13$$

$$Rp\ (\underset{\sim}{\xi}) = C\delta_{po} \qquad 3.14$$

These equations must be solved for $\underset{\sim}{\xi}$ (using the constraints to be λ_p), but are nonlinear. We use the Gauss-Newton iteration method which ignores certain second derivatives to obtain

$$\sum_t\left[2\sum_{k=0}^{N}\frac{\partial F_{n-k}}{\partial \xi_m}\ \frac{\partial F_{n-k}}{\partial \xi_t}+\sum_{p=0}^{s-1}\lambda_p\ \frac{\partial^2 Rp}{\partial \xi_m \partial \xi_t}\right]_i(\xi_t^{i+1}-\xi_t^i)+$$

$$\sum_{p=0}^{s-1}\frac{\partial Rp}{\partial \xi_m}\ (\lambda_p^{i+1}-\lambda_p^i) = -\left[2\sum_{k=0}^{N}(F_{n-k}-e_{n-k})\frac{\partial F_{n-k}}{\partial \xi_m}+\right.$$

$$\left.\sum_{p=0}^{s-1}\lambda_p\ \frac{\partial Rp}{\partial \xi_m}\right]_i \qquad 3.15$$

where i refers to the iteration number.

And the constraints are linearized to

$$\sum_t \frac{\partial Rp}{\partial \xi_t}\ (\xi_t^{i-1}-\xi_t) = -{}^{i}_{Rp} \qquad 3.16$$

These equations may be written in matrix form as:

$$\underset{\approx}{\mathcal{O}}\,\underset{\sim}{\Delta\xi} + \underset{\approx}{E}\,\underset{\sim}{\Delta\lambda} = \underset{\sim}{I} \qquad 3.17$$

$$H\,\underset{\sim}{\Delta\xi} = \underset{\sim}{J} \qquad 3.18$$

where

$$\mathcal{O}_{mt} = 2\sum_{k=0}^{N} \frac{\partial F_{n-k}}{\partial \xi_m}\,\frac{\partial F_{n-k}}{\partial \xi_t} + \sum_{p=0}^{s-1} \lambda_p\,\frac{\partial^2 R_p}{\partial \xi_m \partial \xi_t}$$

$$E_{mp} = \frac{\partial R_p}{\partial \xi_m}$$

$$I_m = -\left[\, 2\sum_{k=0}^{N} (F_{n-k} - e_{n-k})\,\frac{\partial F_{n-k}}{\partial \xi_m} + \sum_{p=0}^{s-1} \lambda_p\,\frac{\partial R_p}{\partial \xi_t}\,\right]$$

$$H_{pt} = \frac{\partial R_p}{\partial \xi_t}$$

$$J_o = -R_p$$

A single matrix equation may be written as:

$$\begin{bmatrix} \underset{\approx}{\mathcal{O}} & \underset{\approx}{E} \\ \underset{\approx}{H} & \emptyset \end{bmatrix} \begin{bmatrix} \underset{\sim}{\Delta\xi} \\ \underset{\sim}{\Delta\lambda} \end{bmatrix} = \begin{bmatrix} \underset{\approx}{I} \\ \underset{\approx}{J} \end{bmatrix} \qquad 3.19$$

This matrix equation is solved. And the results are used to update the iterates

$$\xi^{n+1} = \xi^{n} + \Delta\xi \text{ , etc.} \qquad 3.20$$

Despite the excellent detection/classification capability of the original TRAPS, due to its adaptive nature, it did exhibit signal cancellation phenomena like any adaptive filter (24-26,17), when adapting rapidly. This phenomena is a consequence of the adaptive interaction between the desired signal and interference. This interference could be anything from bandpass noise to a sum of sinusoids suitably spaced in frequency. The existence of these problems in the AARMA model as well as the AR model motivated the development of a more complex first stage of TRAPS. The method developed has similitude with work of Ken Duvall (27).

A schematic of the adaptive noise canceller at its development stage per the semi-annual report is shown in Figure 8. The "primary input" contains a useful signal s, plus interference n_o. The "reference input" contains interference n_1, related to that of the primary input, and is separately obtained in practical systems. The relationship between the two interferences is generally unknown a priori. The adaptive filter has the job of shaping the reference interference to replicate (in the least squared error sense) the primary

interference so that subtraction will remove the interference from the primary input and thereby deliver a much more useful output. It has been shown in (28) that an adaptive filter minimizing output power in the system of Figure 8 causes the system output to be a best least squares estimate of the useful signal s. Unfortunately, under certain circumstances, primary input components may be partially cancelled and distorted despite the fact that they are not correlated with the reference input. This distortion in many cases would be quite acceptable. What is not acceptable are the "non-Weiner" effects due to weight fluctuation which give rise to signal cancellation.

In order to facilitate discussion of the solution to the problem of these non-Weiner effects, consider the example where a sinusoid jammer from a direction other than that of the primary signal, or "off the look direction", causes sinusoidal fluctuations in all the weights. The look direction constraint is unit gain and zero phase over all frequencies. The useful signal arriving in the look direction encounters a unit gain with the main beam so constrained. The sinusoid off the look direction for this example produces a notch along the primary signal flow path to the output by means of the fixed weight filter. Notching phenomena in this system are very much like those of the adaptive canceller of Fig. 8. Signals from the look direction do not appear at the adaptive

filter inputs; only interference is present there. Both signal and jammer are present in the primary signal flow path, and both signal and jammer experience notching at the jammer frequency. With high speed adaptation, there is the danger of eliminating the signal and jammer. The signal cancellation phenomena that has just been discussed is due to interaction between the signal and the jammer in the adaptive beamformer. Since the interaction is the root of the problem, it is useful to consider adaptive structures that separate the signal and the jammer, during the adaptive process. A remedy to the signal cancellation problem will now be discussed. The method devised is based on original work by Ken Duvall (27). The Duvall system incorporates two signal processing systems, one to perform the adaptation, the other to generate the system output signal.

The system incorporated into TRAPS' Stage One is shown in Figure 9. It makes use of two filters per phone. The filter on the right is connected directly to the elements and is used to derive the array output signal. It is, however, a slaved filter rather than the adaptive filter that would usually be expected in this position. The filter on the left is the TRAPS adaptive filter that is connected to the array elements through a subtractive preprocessor. The preprocessor excludes the look-direction signal, but admits jammer signals in a modified form. The adaptive process generates a set of

weights that provides some specified look-direction gain while minimizing (in the least-squares sense) the jammer contribution. These weights are then input into the slaved beamformer to provide the desired signal reception and jammer rejection.

This study of a slaved first stage provided insight into a Stage Two system which not only prevented the signal cancellation, but gave the algorithm perspective as to the path the coefficients were to take, once the error exceeded a specific threshold. This system for Stage Two was labeled "cascade structure" and is detailed in Section 3.3.

3.3 Stage Two Cascade Formalism

What the algorithm detailed in Section 3.2 permits is the obtainment of an accurate ARMA model, which is constrained to remain stable. The accuracy of the ARMA model for the majority of the cases is impressive. Extensive analysis since the semi-annual report showed a weakness in Stage Two of TRAPS, when the noise or signal had a sharp transition in its characteristics. While this weakness was prevalent for only a short span of time and did not deteriorate the model at high SNR, it did reduce localization accuracy at low SNR. This problem was therefore resolved by what we shall term a "cascade structure." The primary goal of this section is to detail the development of a cascade structure to the AARMA model of TRAPS' Stage Two. So that the processing gains allowed by this cascade structure can be best understood, this section will also brief the characteristics of the mean-squared-error surface and the performance of a recursive LMS algorithm for use in the evaluation.

A performance surface in terms of the mean-squared-output error is

$$E(e^2) = E[(y-\hat{y})^2] = E(y^2) + E(\hat{y}^2) - 2E(y\hat{y}) \qquad 3.21$$

So that the association between the cascade structure and the

improvements to the AARMA may be easily understood, the developmental steps will be restricted to the case where the model and the adaptive algorithm are both second order with

$$H(z) = \frac{B(z)}{A(z)} = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2}}{1 - a_1 z^{-1} - a_2 z^{-2}}. \qquad 3.22$$

The assumption that y and $\hat{y}$ are stationary random signals with zero mean and finite variance, and that the input u is white noise with unit power density, allows a normalized error surface to be defined in terms of the transfer functions [29]

$$\mathcal{E} = 1 + \frac{1}{2\pi j \sigma^2} \oint \hat{H}(z)[\hat{H}^*(z) - 2H^*(z)]\frac{dz}{z} \qquad 3.23$$

where

$$\sigma^2 = \frac{1}{2\pi j} \oint [H(z)]^2 \frac{dz}{z} , \qquad 3.24$$

$\hat{H}(z)$ is the transfer function of the adaptive algorithm, and $H^*(z)$ denotes the complex conjugate of $H(z)$.

Attention shall now be confined to the two-pole model in order to illustrate error surface effects with

$$H(z) = \frac{1}{1 - a_1 z^{-1} - a_2 z^{-2}} \qquad 3.25$$

This allows the mapping of the interior of the unit circle in

the z-plane onto a triangular region in the plane associated with the two adaptive coefficients. The interior of the triangle shown in Figure 10 thus represents the stable region for the adaptive coefficients $\hat{a}_1$, $\hat{a}_2$.

Plots illustrating the contours of constant mean-squared-error for three different models are illustrated in Figure 11. These plots emphasize the extreme nonuniformity of the ARMA performance surface. The contours tend to be elongated in shape, so that the gradient vector does not always direct itself toward the minimum where $\mathcal{E} = 0$. A more complete analysis of the error surface of AARMA algorithms is provided in reference [29].

The parameter update formula for an adaptive coefficient vector θ is expressed as

$$\theta_{k+1} = \theta_k + \mu(-\nabla_k) \qquad 3.26$$

where μ is a parameter which controls stability and rate of convergence, and ∇_k is the value (or an estimate) of the error surface gradient where $\theta = \theta_k$. Detail of this can be found in Section 2.1.

For the comparisons to follow it is important to realize that for a fixed point on the mean-squared-error surface, the estimate obtained from the recursive LMS algorithm is unbiased. To see this, we define a coefficient vector

$$\hat{\theta} = [\hat{a}_1 \ \hat{a}_2 \ \hat{b}_0 \ \hat{b}_1 \ \hat{b}_2]^T \ . \qquad 3.27$$

The true gradient vector is therefore expressed as

$$\nabla_k = \frac{\partial \xi_k}{\partial \hat{\theta}} \qquad 3.28$$

where

$$\xi_k = E(e_k^2) \ . \qquad 3.29$$

The gradient estimate for the recursive LMS algorithm is derived in terms of the instantaneous squared output error as

$$\hat{\nabla}_k = \frac{\partial \hat{\xi}_k}{\partial \hat{\theta}} \qquad 3.30$$

where

$$\hat{\xi}_k = e_k^2 \ . \qquad 3.31$$

We now define the expected value of the gradient estimate over k as the average of an arbitrarily large number of samples:

$$E(\hat{\nabla}_k) = \frac{1}{N} \sum_{k=1}^{N} \hat{\nabla}_k \qquad 3.32a$$

$$= \frac{1}{N} \sum_{k=1}^{N} \frac{\partial e_k^2}{\partial \hat{\theta}} \qquad 3.32b$$

$$= \frac{\partial}{\partial \hat{\theta}} \left[\frac{1}{N} \sum_{k=1}^{N} e_k^2 \right] \qquad 3.32c$$

$$= \frac{\partial}{\partial \hat{\theta}} \left[E(e_k^2) \right] \qquad 3.32d$$

$$= \frac{\partial \xi_k}{\partial \theta} \triangleq \nabla_k \quad . \qquad 3.32e$$

The fact that $E(\hat{\nabla}_k) = \nabla_k$ implies that the gradient estimate is asymptotoically unbiased.

Simulation Results:

Again we will restrict our attention to the two-pole model and fix the numerator coefficients: $b_0 = \hat{b}_0 = 1$, $b_1 = b_2 = \hat{b}_1 = \hat{b}_2 = 0$. In Figure 12, parameter track plots illustrate the performance of the recursive LMS algorithm for each of three error surfaces.

It is apparent that the gradient estimates are very noisy in these examples. In a nonideal environment, where additional noise is introduced by signal measurement and other uncontrollable sources, generally this amount of estimate noise would be unacceptable. Smoother parameter tracks may be obtained at the expense of increased convergence time by reducing the size of n. This is illustrated in Figure 13. The recursive LMS algorithm did converge in every example when p was

properly selected. When some a priori knowledge of the model characteristics is not available, choice of an appropriate value remains an empirical process.

From the preceeding examples, we see that a constant value of n does not work well for the extremely nonuniform error surfaces. Ideally, we would like to take small steps when the gradient is very steep to avoid bouncing from side to side across the "valley." If n is so selected, when the coefficients reach the floor of the valley, convergence to the minimum will be very sluggish due to the small gradient [29].

The definition of a cascade structure can now be attempted. The basic procedure for Stage Two will still follow the technique described in Section 3.2. When e reaches a predetermined upper limit, the algorithm switches to a cascade structure. When the cascade structure is in position the transfer functions are of the form

$$H(z) = \prod_{j=1}^{N} Hj(z) \qquad 3.33$$

where

$$H_j(z) = \frac{B_j(z)}{A_j(z)} = \frac{1 + b_{1j}z^{-1} + b_{2j}z^{-2}}{1 - a_{1j}z^{-1} - a_{2j}z^{-2}} \quad j < N \qquad 3.34$$

$$H_N(z) = \frac{B_N(z)}{A_N(z)} = \frac{b_0 + b_{1N}z^{-1} + b_{2N}z^{-2}}{1 - a_{1N}z^{-1} - a_{2N}z^{-2}} \; . \qquad 3.35$$

The adaptive cascade structure is illustrated in Figure 14.

Applying the steepest descent criteria defined earlier, the coefficient updates are obtained of the form

$$\hat{b}_{ij}(k+1) = \hat{b}_{ij}(k) + \mu\left(-\frac{\partial \overline{e^2}}{\partial \hat{b}_{ij}}\right)\begin{cases} i = 1, 2 & j = 1, 2 \ldots N-1 \\ i = 0, 1, 2 & j = N \end{cases} \qquad 3.36a$$

$$\hat{a}_{ij}(k+1) = \hat{a}_{ij}(k) + \eta\left(-\frac{\partial \overline{e^2}}{\partial \hat{a}_{ij}}\right) \quad i = 1, 2 \qquad j = 1, 2 \ldots N \qquad 3.36b$$

As in the derivation of the recursive LMS algorithm, we now assume that the instantaneous output error is a local estimate for the mean-squared error to obtain

$$\frac{\partial \overline{e^2}}{\partial \hat{b}_{ij}} \approx 2e(k)\frac{\partial e(k)}{\partial \hat{b}_{ij}} \qquad 3.37a$$

$$\frac{\partial \overline{e^2}}{\partial \hat{a}_{ij}} \approx 2e(k)\frac{\partial e(k)}{\partial \hat{a}_{ij}} \qquad 3.37b$$

Now we make use of the fact that the model output y is independent of the adaptive coefficients and define the partial derivatives

$$\beta^N_{ij}(k) \triangleq \frac{\partial \hat{y}_N(k)}{\partial \hat{b}_{ij}} \qquad 3.38a$$

$$\alpha_{ij}^{N}(k) \triangleq \frac{\partial \hat{y}_N(k)}{\partial \hat{a}_{ij}} \quad . \qquad 3.38b$$

Combining Equations 3.36 - 3.38, the coefficient updates are obtained as follows

$$\hat{b}_{ij}(k+1) = \hat{b}_{ij}(k) + \mu e(k)\beta_{ij}^{N}(k) \begin{cases} i = 1, 2 & j = 1, 2 \ldots N-1 \\ i = 0, 1, 2 & j = N \end{cases} \qquad 3.39a$$

$$\hat{a}_{ij}(k+1) = \hat{a}_{ij}(k) + \eta e(k)\alpha_{ij}^{N}(k) \quad i = 1, 2 \quad j = 1, 2 \ldots N. \qquad 3.39b$$

In order to derive expressions for α_{ij}^{N} and β_{ij}^{N} defined in Equation 3.38, we will make use of a series of intermediate values

$$\beta_{ij}^{n}(k) \triangleq \frac{\partial \hat{y}_n(k)}{\partial \hat{b}_{ij}} \quad n = 1, 2 \ldots N \qquad 3.40a$$

$$\alpha_{ij}^{n}(k) \triangleq \frac{\partial \hat{y}_n(k)}{\partial \hat{a}_{ij}} \quad n = 1, 2 \ldots N \quad . \qquad 3.40b$$

Since the output of stages 1, 2 ... j - 1 is unaffected by the coefficients of stage j, we see that

$$\beta_{ij}^{n}(k) = 0 \quad n < j \qquad 3.41a$$

$$\alpha_{ij}^{n}(k) = 0 \quad n < j \quad . \qquad 3.41b$$

When n = j, the values are obtained exactly as for the second order section in Section 2.1,3.3 with the input now equal to $\hat{y}_{j-1}$ (or u(k) if j=1) and output $\hat{y}_j$. Suppressing the time dependence of the adaptive coefficients for notational convenience we obtain

$$\beta_{ij}^{j}(k) = \hat{y}_{j-1}(k-i) + \hat{a}_{1j}\beta_{ij}^{j}(k-1) + \hat{a}_{2j}\beta_{ij}^{j}(k-2)$$

$$\alpha_{ij}^{j}(k) = \hat{y}_{j}(k-i) + \hat{a}_{1j}\alpha_{ij}^{j}(k-1) + \hat{a}_{2j}\alpha_{ij}^{j}(k-2) \quad . \qquad 3.42b$$

For n > j, the partial derivatives are formulated as filtered versions of β_{ij}^{j} and α_{ij}^{j}:

$$\beta_{ij}^{n}(k) = \hat{b}_{0n}\beta_{ij}^{n-1}(k) + \hat{b}_{1n}\beta_{ij}^{n-1}(k-1) + \hat{b}_{2n}\beta_{ij}^{n-1}(k-2)$$
$$+ \hat{a}_{1n}\beta_{ij}^{n}(k-1) + \hat{a}_{2n}\beta_{ij}^{n}\beta(k-2) \qquad 3.43a$$

$$\alpha_{ij}^{n}(k) = \hat{b}_{0n}\alpha_{ij}^{n-1}(k) + \hat{b}_{1n}\alpha_{ij}^{n-1}(k-1) + \hat{b}_{2n}\alpha_{ij}^{n-1}(k-2)$$
$$+ \hat{a}_{1n}\alpha_{ij}^{n}(k-1) + \hat{a}_{2n}\alpha_{ij}^{n}(k-2) \quad . \qquad 3.43b$$

Note that

$$\hat{b}_{0n} = 1 \qquad n < N \quad . \qquad 3.43c$$

If we now assume that small changes are made at each iterative cycle so that $\hat{H}_k(z) \simeq \hat{H}_{k+1}(z)$, we obtain a slightly simpler version

$$\hat{b}_{ij}(k+1) = \hat{b}_{ij}(k) + \mu e(k) \beta_j^N(k-i) \begin{cases} i = 1, 2 & j = 1, 2 \ldots N-1 \\ i = 0, 1, 2 & j = N \end{cases} \quad 3.44a$$

$$\hat{a}_{ij}(k+1) = \hat{a}_{ij}(k) + \eta e(k) \alpha_j^N(k-i) \quad i = 1, 2 \quad j = 1, 2 \ldots N-1 \quad 3.44b$$

where β_j^N and α_j^N are formulated as illustrated in Figure 15. An example illustrating the convergence of a four pole system is provided in Figure 16.

3.4 Evaluation of Cascaded AARMA on Simple Simulations

A small class of AARMA algorithms with comparable computational complexity and proven convergence characteristics will be used to demonstrate the capability of the cascade structure. By choosing the order of the adaptive algorithm to be at least as high as that of the model, a unimodal performance surface is assured [29].

In addition to the adaptive algorithms presented earlier, we will employ the SHARF algorithm. The analysis by Ljung [30-31] provides a convergence proof for the RLMS algorithm. Since SHARF is not a gradient search algorithm, the convergence proof relies upon the concept of hyperstability [32-36].

A direct form realization of the transfer function

$$\hat{H}(z) = \frac{\hat{B}(z)}{\hat{A}(z)} = \frac{\hat{B}_0 + \hat{b}_1 z^{-1} + \ldots + \hat{b}_N z^{-N}}{1 - \hat{a}_1 z^{-1} - \ldots - \hat{a}_N z^{-N}} \qquad 3.45$$

is utilized by both the RLMS and SHARF algorithms.

Since the second order RLMS algorithm discussed in the previous Section was constrained to remain stable during the adaptive process, use will also be made of the general Nth order unconstrained realization. This then provides a wide range of algorithms for the following examples. The adaptive updates are given by

$$\hat{b}_i(k+1) = \hat{b}_i(k) + \mu e(k)\beta(k-i) \quad i = 0, 1 \ldots N \quad 3.46a$$

$$a_i(k+1) = a_i(k) + \eta e(k)\alpha(k-i) \quad i = 1, 2 \ldots N \quad 3.46b$$

where α, β and e are formulated as in Figure 15.

Use of the SHARF algorithm requires the selection of a set of error smoothing coefficients, c_1, c_2 ... c_p [32]. In order to satisfy the hyperstability criterion, these parameters must be chosen so that

$$\mathrm{Re}\left[\frac{C(z)}{A(z)}\right] > 0 \quad 3.47$$

where

$$C(z) = 1 + c_1 z^{-1} + c_p z^{-p} \quad 3.48$$

and A(z) is the denominator of the fixed transfer function H(z). As a result, some a priori knowledge of the model is required. An alternate technique has been proposed which allows recursive estimation of the C(z) coefficients [37]. Here we will use the fixed C(z) polynomial. The adaptive updates for the SHARF algorithm may be expressed as

$$\hat{b}_i(k+1) = \hat{b}_i(k) + \mu u(k-i)\, v(k) \quad i = 0, 1 \ldots N \quad 3.49a$$

$$\hat{a}_i(k+1) = \hat{a}_i(k) + \pi\hat{y}(k-i)\, v(k) \quad i = 1, 2 \ldots N \qquad 3.49b$$

where u, $\hat{y}$ and v are formulated as in Figure 15 [33].

A second order model with transfer function

$$H(z) = \frac{.2871 - .4466z^{-1}}{1. - 1.2971z^{-1} + .6949z^{-2}} \qquad 3.50$$

will be the first example to be used for comparison. The plots in Figures 17 and 18 illustrate the reduction in output error versus the number of data samples used for four different cases.

In Figure 17.a, the RLMS algorithm defined by 3.46 was used. With convergence factors $\pi = \mu = .07$, the resultant steady-state transfer function was

$$\hat{H}(z)_{RLMS} = \frac{.2871 - .4466z^{-1} - .0001z^{-2}}{1. - 1.2970z^{-1} + .6949z^{-2}} \qquad 3.51$$

Figure 17.b illustrates the performance of Cascade AARMA (CAARMA) algorithm defined by (27). Using convergence factors $\pi = \mu = .5$, the steady-state transfer function was

$$\hat{H}(z)_{CAARMA} = \frac{.2871 - .4466z^{-1}}{1. - 1.2971z^{-1} + .6949z^{-2}} \qquad 3.52$$

In Figure 18, the SHARF algorithm defined in (3.49) was used. For the transfer function defined by (3.50), we can show that

$$\text{Re}\left[\frac{1}{A(z)}\right] > 0 \quad . \tag{3.53}$$

Thus, C(z) = 1 so that no error smoothing is performed. Figure 18.a illustrates the performance of SHARF for this case with convergence factors $\eta = \mu = .07$. The steady-state transfer function was

$$\hat{H}(z)_{SHARF} = \frac{.2871 - .4466z^{-1} - .0001z^{-2}}{1. - 1.2970z^{-1} + .6947z^{-2}} \tag{3.54}$$

It has been shown that choice of the C(z) coefficients also affects the rate of parameter convergence [32]. This is illustrated by Figure 18.b. Here we let

$$C(z) = 1. - 1.25z^{-1} \tag{3.55}$$

and again use convergence factors $\eta = \mu = .07$. The resultant steady-state transfer function was

$$H(z)_{SHARF} = \frac{.2871 - .4466z^{-1}}{1. - 1.2971z^{-1} + .6949z^{-2}} \tag{3.56}$$

Note the improvement in the rate of parameter convergence which results from the error smoothing. For further analysis of the relationship between the C(z) polynomial and the rate of convergence, see References [32-36].

From the results of these simulations, we see that the CAARMA algorithm evidenced the best performance in terms of the number of data samples required for coefficient convergence. An attempt was made to maximize the rate of convergence in all cases by adjusting η and μ, but as was remarked previously, SHARF convergence characteristics may be improved by a different selection of C(z).

It has been shown that when the order of the adaptive algorithm exceeds that of the model and the performance surface is unimodal, the minima is distributed [29]. A fourth order adaptive algorithm will now be used as an example in an attempt to identify a third order model with transfer function

$$H(z) = \frac{.0563(1. + z^{-1})(1. - 1.0166z^{-1} + z^{-2})}{(1. - .6830z^{-1})(1. - 1.4461z^{-1} + .7957z^{-2})} \qquad 3.57$$

This is the digital equivalent of a low-pass elliptic filter which was obtained via the bilinear transformation [38]. The plots in Figures 19 and 20 illustrate the reduction in output error versus the number of data samples for four different adaptive algorithms.

The cascade structure algorithms were used in Figure 19. Figure 19.a illustrates the performance of the CAARMA algorithm defined by (3.42). With convergence factors $\eta = \mu = .025$, the steady-state transfer function was

$$\hat{H}(z)_{CAARMA} = \left(\frac{1. - 1.0164z^{-1} + .9999z^{-2}}{1. - 1.4462z^{-1} + .7958z^{-2}} \right)$$

$$\left(\frac{.0562 + .0769z^{-1} + .0207z^{-2}}{1. - .3121z^{-1} - .2534z^{-2}} \right) . \qquad 3.58$$

The second stage can be factored as

$$\frac{.0562(1. + .998z^{-1})(1 + .370z^{-1})}{(1 - .683z^{-1})(1 + .371z^{-1})} , \qquad 3.59$$

thus illustrating the existence of a common factor in $\hat{H}(z)$ which results from the distributed minima.

The two direct form adaptive algorithms defined at the beginning of this Section were employed during the simulation illustrated in Figure 20. The direct form equivalent of the model defined by (3.57) is

$$H(z) = \frac{.0563(1. - .0166z^{-1} - .0166z^{-2} + z^{-3})}{1. - 2.1291z^{-1} + 1.7834z^{-2} - .5435z^{-3}} \qquad 3.60$$

Figure 20.a was obtained via the RLMS algorithm defined by (3.46) with convergence factors $\mu = .01$, $\eta = .005$. The resultant steady-state transfer function can be expressed as

$$\hat{H}(z)_{RLMS} = \frac{.0562(1. + .867z^{-1})(1. - .0165z^{-1} - .0165z^{-2} + 1.002z^{-3}}{(1. + .867z^{-1})(1. - 2.130z^{-1} + 1.7836z^{-2} - .5432z^{-3})}$$

3.61

thus demonstrating the existence of a common factor due to the distrubuted minima. A fourth order SHARF algorithm was employed during the simulation illustrated in Figure 20.b. To ensure satisfaction of the strictly positive real condition, we let

$$C(z) = 1. - 1.4461z^{-1} + .7957z^{-2} \qquad 3.62$$

to obtain

$$\mathrm{Re}\left[\frac{C(z)}{A(z)}\right] = \mathrm{Re}\left[\frac{1}{1. - .6830z^{-1}}\right] > 0 \; . \qquad 3.63$$

With convergence factors $\mu = \eta = .07$, the resultant steady-state transfer function was

$$\hat{H}(z)_{SHARF} = \frac{.062(1. + .762z^{-1})(1. - .0164z^{-1} - .0165z^{-2} + .9998z^{-3})}{(1. + .762z^{-1})(1. - 2.128z^{-1} + 1.7836z^{-2} - .5434z^{-3})} \qquad 3.64$$

This example serves a dual purpose. It illustrates the performance of the multistage cascade algorithms and demonstrates that when the performance surface minima is distributed, the transfer function $\hat{H}(z)$ contains a common factor.

It is evident from the illustrations in Figures 19 and 20 that the cascade structure algorithms yielded the

superior rate of convergence under the conditions previously described. Maintaining stability during the adaptive process was found to be the primary limiting factor for both RLMS and SHARF. It is possible that a different selection of C(z) could improve the rate of convergence of SHARF, however the choice given by (3.62) already assumes a great deal of a priori knowledge of model parameters.

An idealized operating environment has been used in the previous simulations, so as to yield insight into the characteristics of this class of adaptive algorithms. The performance of the adaptive algorithms presented earlier will now be examined under nonideal operating conditions.

The first condition that will be examined is that of the algorithm response to insufficient model order. Guaranteed convergence to the global minima requires that the performance surface be unimodal when gradient search techniques are employed. It has been shown that when the order of the adaptive algorithm is less than that of the model, the performance surface may contain multiple minima even though the system identification configuration is utilized and the input is white noise [29]. In this case, steady-state adaptive coefficients are dependent upon their initial values [30].

Since the performance surface is quadratic with respect to the numerator coefficients, pole adaptation is our primary concern. Here we will assume a sixth order all-pole

model with transfer function

$$H(z) = \frac{.1}{(1. - 1.2z^{-1} + .6z^{-2})(1. + 1.2z^{-1} + .6z^{-2})(1. - .64z^{-2})} \qquad 3.65$$

The corresponding pole locations are symmetrically distributed in the z-plane as illustrated in Figure 21.

Simulations were performed using a fourth order CAARMA algorithm with three different sets of initial conditions on the stage two coefficients. Zero initial conditions were used for stage one coefficients. The resultant steady-state pole locations are illustrated in Figure 22. Plots demonstrating the reduction in output error for each case are provided in Figure 23. Note that the error reduction in Figure 23.c was minimal, thus clearly illustrating the primary hazard of multimodal performance surfaces.

The next example that shall be considered is that of delay adjustment. Earlier it was pointed out that even when the order of the adaptive algorithm is sufficient, the cascade structure defined by Figure 14 may not allow the unique solution $\hat{H}(z) = H(z)$ because of a delay mismatch. By using the alternate structure illustrated in Figure 24, however, this problem can be alleviated.

Consider a fourth order model with transfer function

$$H(z) = \frac{.5z^{-4}}{(1. - 1.2z^{-1} + 6z^{-2})(1 - z^{-1} + .4z^{-2})} \qquad 3.66$$

Utilizing a fourth order cascade structure, we see that

$$\hat{B}(z) = (1 + \hat{b}_{11}z^{-1} + \hat{b}_{21}z^{-1})(\hat{b}_{02} + \hat{b}_{12}z^{-1} + \hat{b}_{22}z^{-2})$$

$$= \hat{b}_{02} + (\hat{b}_{11}\hat{b}_{02} + \hat{b}_{12})z^{-1} + (\hat{b}_{21}\hat{b}_{02} + \hat{b}_{11}\hat{b}_{12} + \hat{b}_{22})z^{-2} + (\hat{b}_{21}\hat{b}_{12} + \hat{b}_{11}\hat{b}_{22})z^{-3} + (\hat{b}_{21}\hat{b}_{22})z^{-4} \qquad 3.67$$

$$3.68$$

where at lease one of the stage two $\hat{b}$ coefficients must be non-zero. The $\hat{B}(z)$ polynomial defined by (3.68) cannot match the numerator of the transfer function in (3.66). If we now incorporate ℓ units of pre-delay as illustrated in Figure 24, we obtain the adaptive numerator polynomial

$$z^{-\ell}\hat{B}(z) = \hat{b}_{02}z^{-\ell} + (\hat{b}_{11}\hat{b}_{02} + \hat{b}_{12})z^{-\ell-1} + (\hat{b}_{21}\hat{b}_{02} + \hat{b}_{11}\hat{b}_{11} + \hat{b}_{22})z^{-\ell-2} + (\hat{b}_{21}\hat{b}_{12} + \hat{b}_{11}\hat{b}_{22})z^{-\ell-3} + (\hat{b}_{21}\hat{b}_{22})z^{-\ell-4}$$

$$3.69$$

From (3.69) we see that we can now match the model B(z) in three cases:

$$z^{-2}\hat{B}(z) = B(z) \; ; \quad \hat{b}_{22} = .5 \qquad 3.70a$$

$$z^{-3}\hat{B}(z) = B(z) \; ; \quad \hat{b}_{12} = .5 \qquad 3.70b$$

$$z^{-4}\hat{B}(z) = B(z) \; ; \quad \hat{b}_{02} - .5 \qquad 3.70c$$

where all unspecified $\hat{b}_{ij}$ coefficients are zero.

Simulations were performed for two cases using the fourth order CAARMA algorithm with $\mu = \eta = .5$. Figure 25 illustrates the results obtained with four units of pre-delay. The steady state transfer function was

$$\hat{H}(z) = \frac{(1. - .003z^{-1} + .000z^{-2})(.500 + .0041z^{-1} - .004z^{-2})}{(1. - 1.199z^{-1} + .600z^{-2})(1. - 1.008z^{-1} + .401z^{-2})} \qquad 3.71$$

The resultant numerator polynomial can be expressed as

$$z^{-4}\hat{B}(z) = .500z^{-4} + .002z^{-5} - .004z^{-6} \; . \qquad 3.72$$

Two units of pre-delay were used during the simulation illustrated in Figure 25.b. The steady-state transfer function was

$$\hat{H}(z) = \frac{(1. - .005z^{-1} - .006z^{-2})(.000 - .001z^{-1} + .500z^{-2})}{(1. - 1.198z^{-1} + .600z^{-2})(1 - 1.006z^{-1} + .401z^{-2})} \qquad 3.73$$

which yields the numerator polynomial

$$z^{-2}\hat{B}(z) = .001z^{-3} + .500z^{-4} - .002z^{-5} - .003z^{-6} \; . \qquad 3.74$$

Further analysis shows that as long as the choice of ℓ is no more than two units less than the number of units of pure delay in the model, the cascade structure of Figure 24 allows the solution $z^{-\ell}\hat{H}(z) = H(z)$ regardless of system order.

3.5 Parameters of TRAPS' which Provide Enhanced Estimation of Seismic Noise Structure, Using Arrays

The frequency-wavenumber (f-k) power spectral density of a seismic noise field is one common way to summarize the velocity and frequency properties of the noise field (39). The determination of the f-k structure of seismic waves is one of the most powerful techniques available for the study of the mode structure and the properties of these waves. Studies of the f-k spectral density made available by array data has given insight into the problem (40-45). This section focuses on TRAPS' Stage One noise spectrum and the f-k delineation made possible by time tuning specific parameters of the CAARMA model of the preceeding Sections.

As cited in Section 2.1 and 3.2 the definition of noise power involves a number of assumptions concerning the time and space stationarity of the seismic noise field. Also, each event is different to some extent and a judgment concerning the validity, over temporal and spatial regions of interest, of such assumptions must be made. One of the advantages of TRAPS is that due to the adaptive nature of the processing system most assumptions are reduced or consequential biases are eliminated. One example is that the assumption of stationarity can be reduced to quasi-stationarity. Also, Stage

One places on each sensor in an array an independent adaptive filter. These filters converge to those data values, which were made local to that sensor. While this technique has far reaching effects in yielding an increased SNR, this independent adaptation allows automatic removal of some effects of seismometer gain variations. For the filters acting on the individual geophones of the array a sufficiently long adaptation time was selected in order to accomodate the variations in ambient noise structure without disturbing the transient signal signature. For the filters acting on the beamformed data, the adaptation time varied from one event to another in order to incorporate the local variance of each event. Furthermore, in many processing techniques a too small of a window or model order a bias is injected into the scheme. If the length is too small the noise will appear consistently less spatially organized than it really is. While the order of the CAARMA model is critical, this bias is bypassed by TRAPS. This is due to the fact that the adaptation gives the CAARMA model the capability of a model with much greater order (46). The adaptive nature of TRAPS also eliminates most problems associated with seasonal changes.

3.6 Technical Results of CAARMA model on NORESS Data

The adaptive ARMA model which is the foundation of the CAARMA model and the slaved first Stage which prohibits the effects of jamming (nearby man-made disturbances, local seismic disturbances, etc.) and signal cancellation each attacked a different weakness in the original TRAPS. Figures 26 through 28 of this report are provided to allow insight into the gain which each of these two improvements provided. The reader is referred to the preceeding semi-annual report if it is desired to obtain a separate detailed evaluation of each improvement, since only the results of the complete CAARMA model shall be presented at this time.

The complete CAARMA model processing system enables TRAPS to detect, locate and classify seismic events for nuclear monitoring purposes by providing the necessary information for accurate results. Since the Technical Results of Section 4 shall be demonstrated by means of TRAPS' location capability, this Section shall present the results of this Section by way of instantaneous power spectral plots. Figure 29 is a raw data plot of a NORESS event. The event detail is listed on the Figure. Further event information can also be found in Table 2. Figure 30.a is a plot of the instantaneous power spectrum of the event represented in Figure 29, using the parameters indicated by the statistical review of the data to

be optimum for event detection. Figures 30.b-30.f are provided to clarify how various CAARMA parameters can be used to pull out specific features of an event. Since the results on Section 4 shall be presented in a manner to afford proper evaluation and comparison of the final TRAPS, Figure 30 is provided so that visual study of the CAARMA model can occur in a familiar mode, the spectrum.

4.0 TASK 3: DETECTION, LOCATION AND CHARACTERIZATION OF NORESS EVENTS BY TRAPS

4.1 Introduction

The purpose of Task 3 for the contract was to characterize the class of transients which compose the signal, so that automatic identification can accurately occur. The frequency-wave vector power spectrum describes the distribution of power versus vector and temporal frequency, or equivalently, versus propagation direction and temporal frequency. The frequency wave vector spectrum is important, since the detection and bearing estimation of the source is based upon information contained in the power spectral estimate. Thus among the generic signal processing concerns of an array system is the estimation of the frequency-wavenumber spectrum of the signal and ambient noise field. The knowledge of these spectrums is essential for the design of effective detection and estimation algorithms, when the received signal is contaminated by ambient noise.

In order to demonstrate the logic for the direction taken by the Task 3 research, the equations and discussion of Section 3 must be called to attention. It follows from these equations that the CAARMA coefficients and poles and zeros completely determine the frequency-wavenumber spectrum. The

event is thus represented, not by a set of discrete values of the spectrum or its integral, but by a number of parameters. These parameters being the poles and zeros of a rational function and the coefficients of the CAARMA model. This Section demonstrates how this information is utilized to yield a highly accurate localization algorithm for both near (events in Sweden and Norway) and far (Russian events) field applications. Section 4.2 presents a method for improved time delay estimation using the poles of a CAARMA model. Section 4.3 discusses an innovative scheme for enhanced detection, location and characterization made possible by the coefficients and residual of the CAARMA model. Section 4.4 shows the results of these methods on NORESS/NORSAR data.

4.2 An Improved Time Delay Estimation for Stage Three of TRAPS by means of Pole/Zero Decomposition

The semi-annual report described a localization algorithm that was quite successful for near field cases. Unfortunately, the time delay portion of the algorithm showed a decline in performance for such far field locations as Russia. The semi-annual report demonstrated how the poles of an AR process, due to their signal structure sensitivity, can be used to detect a wave front. This section shall expand on this concept to illustrate how the poles and zeros of an AARMA process can be used to accurately determine the time differences of arrivals; even for far field applications.

A variety of time delay estimation techniques have been documented which adequately address the problem of time delay estimation [47]. They can basically be classified into two categories. One is the crosscorrelation method, which was the method originally used in TRAPS [48]. The second mode of time delay models the time delay as a finite impulse response filter with one process, say x(t), as the input. This maintains similitude to an MA filter. The time delay estimate is extracted from the filter coefficients. The innovational procedure for time delay estimation to now be described is one that is based on the adaptive ARMA developed under this contract and detailed in earlier sections.

Before the correlation is established between the adaptive ARMA model and the time delay analysis, a few foundational statements. A signal is assumed at all geophones within some interval of time (T_1, T_2). (In the following Section it shall be demonstrated that this assumption is not required for this new technique to work, but we shall maintain the assumption at this time in order to facilitate discussion.) Let $y_o(t)$ be the filtered time series measured at some reference geophone. Let $y_1(t)$ denote the filtered series on the second phone. Then

$$y_o(t) = s(t) + n_o(t) \qquad 4.1$$

and

$$y_1(t) = s(t-\tau) + n_1(t) \qquad 4.2$$

where t (possibly negative) is the time by which the received signal at geophone 1 lags the received signal at geophone o and s(t) denotes the unknown filtered signal waveform. We assume only that this "signal" is within the interval. $n_o(t)$ and $n_1(t)$ are assumed to be stationary, uncorrelated, zero mean Gaussian processes with variances $\sigma^2_{n_o}$ and $\sigma^2_{n_1}$, respectively.

The time argument at which the maximum of the function

$$h(t) = \int_{-\infty}^{\infty} \frac{S_{y_o y_1}}{S_{y_o y_o}} (\omega) e^{j\omega t} \, d\omega \qquad 4.3$$

occurs is equal to τ. This estimator is equivalent to the Roth processor [49] for time delay estimation, although the attainment is different. In equation (4.3) $S_{y_o y_1}(\omega)$ is the cross-power spectrum between $y_o(t)$ and $y_1(t)$ and $y_o y_o$ is the power spectra of $y_o(t)$. In order to solve τ in ARMA terms let it be noted that [50]

$$\frac{S_{y_o y_1}}{S_{y_o y_o}} (\omega) = \frac{\sum_{k=-l}^{l} \hat{a}_k e^{-j\omega k}}{\sum_{k=0}^{p} \hat{b}_k e^{-j\omega k}}, \quad \hat{b}_o = 1 \qquad 4.4$$

Thus the time delay estimation is a product of the ARMA process. This method was shown to be not only feasible, but relatively superior to most methods, as long as the SNR was greater than OdB.

Referencing our earlier discussion on the adaptive ARMA model of TRAPS and the section that follows, the time delay problem would be better resolved by using only the source poles of an ARMA model, since the zeros are physical related to the noise. Also means must be available for multiple source time delay estimation, due to multipath. In 1979 a new system identification approach for time delay estimation was suggested by Morf [51], using a multichannel rational model for the array signals. This model, which is useful for random sources with various spectra, was later developed by Porat and

Friedlander [52] for the purpose of time delay estimation of unknown sources. The theory was based on an ARMA model. Unfortunately not only did amplitude attenuation need to be calculated, but the receiver noises were assumed to be white and spatially uncorrelated.

The model we propose to use in TRAPS for time delay estimation, based on the adaptive ARMA model detailed earlier, not only produces the time delay when the noise may be correlated among the sensors and in time, but is also capable of calculating fractional delays. While it shall be demonstrated that this technique is not only highly accurate, it must be noted that an advantage of the technique to be discussed lies in its computational simplicity. Compared to the generalized cross-correlation technique, the new time delay technique does not require prior knowledge of the signal and noise spectral densities. The new method also has the advantage that it is completely independent of the attenuation factors.

We now expand on equation (4.3) to give the cross spectrum (or the z-transform of the cross correlation) between y_o and y_1 in a multiple source environment as given by

$$S_{y_o y_1}(z) = G_{y_o k} G_{y_1 k} z^{-\Delta y_o y_1 k} \frac{b_k(z) b_k^*(1/z^*)}{a_k(z) a_k^*(1/z^*)} + S^n_{y_o y_1}(z) \qquad 4.5$$

where $\Delta y_o y_1 k = \tau y_o k - \tau y_1 k$ is the desired time delay of

arrival of a signal emitted from the K^{th} source to the receivers y_o and y_1, and $S^n_{y_o y_1}(z)$ represents the possible correlation between noise components $n_{y_o}(t\ T_s)$ and $n_{y_1}(t\ T_s)$. Here t denotes the time index, while T_s denotes the sampling interval. Before addressing the actual time delay problem it is useful to expand each of the source transfer functions into the partial fraction decomposition

$$\frac{b_k(z)}{a_k(z)} = \sum_{r=1}^{q_k} \frac{C_{kr}}{1-\lambda_{kr} z^{-1}} \qquad 4.6$$

where $\{Ckr\}$ are the residues of $\{b_k(z)/a_k(z)\}$ at the corresponding poles $\{\lambda_{kr}\}$. Substituting 4.6 into 4.5 now gives

$$S_{y_o y_1}(z) = \sum^{m} G_{y_{ok}} G_{y_{1k}} z^{-\Delta y_o y_{1k}} \sum_{r=1}^{q_k} \sum_{s=1}^{q_k} \frac{Ckr}{1-\lambda_{kr} z^{-1}} \frac{C^*ks}{1-\lambda^*_{ks}} + S^n_{y_o y_1}(z) \qquad 4.7$$

Multiplying both the numerator and denominator by $1 - \lambda_{kl} z^{-1}$ and considering the desired limit the following is obtained

$$\lim_{z \to \lambda_{kl}} \frac{S_{y_o y_1}(z)}{S^*_{y_o y_1}(1/z^*)} = \lambda_{kl}^{-2\Delta y_o y_1 k} \qquad 4.8$$

While the applications of these equations appear conditional,

the source transfer functions may have repeated poles and/or be nonproper. This can be proven using the Laurent expansion [53] of $S_{y_0 y_1}(z)$ near each pole by L'Hopital's rule. The procedure for the time delay estimation results from equation 4.8. From the phase part of equation 4.8 results the following time delay estimation

$$\hat{\Delta}_{y_0 y_{1k}} = \frac{\phi\{\hat{S}^{*}_{y_0 y_1}(1/\hat{\lambda}^{*}_{k1})\} - \phi\{\hat{S}_{y_0 y_1}(\hat{\lambda}_{k1})\} + 2t\pi}{2\phi\{\hat{\lambda}_{k1}\}} \qquad 4.9$$

where

$$\phi(x) = tg^{-1}\left(\frac{Im\{x\}}{Re\{x\}}\right) \qquad 4.10$$

The ambiguity term in 4.9 is resolved by the amplitude part of 4.8 which is

$$\hat{\Delta}_{y_0 y_{1k}} = \frac{\log|\hat{S}^{*}_{y_0 y_1}(1/\hat{\lambda}^{*}_{k1})| - \log|\hat{S}_{y_0 y_1}(\hat{\lambda}_{k1})|}{2\log|\hat{\lambda}_{k1}|} \qquad 4.11$$

4.3 Innovative Acoustic Signature Identification by Means of CAARMA Model Output for Accurate Location and Classification

The purpose of this Section is to demonstrate how the information made available by the CAARMA model of TRAPS' Stage One and Two is employed to yield a highly accurate localization and classification algorithm for both near and far field applications. The procedure begins by first isolating a seismic event, using a Weighted Prediction Residual (WPR). This WPR, which will be presented in Section 4.3.1, is then used to elicit the signature of a particular event with the aid of a segmentation procedure. This WPR segmentation technique is analyzed in Section 4.3.2. The WPR views the waveform from a statistical perspective, while the segmentation permits the refinement of the represented waveform a structural aspect. This WPR segmentation procedure is able to classify an event as an earthquake or explosion, use large phase irregularities that categorize specific field data to improve localization accuracy, as well as prepare the weighted CAARMA coefficients (WPC) that are mathematically presented in Section 4.3.1 and time delay information provided by the method of Section 4.2 for entry into a feedback system with a maximum likelihood scheme for an extremely efficacious localization algorithm that will be detailed in Section 4.3.3.

4.3.1 Weighted Prediction Residuals (WPR) and Coefficients (WPC) Based on a CAARMA Model to Determine Statistical Waveform Signature

A valid supposition is that such statistical approaches as Residual (LPR), Linear Prediction Coefficients (LPC), poles and zeros, etc. are useful tools to discriminate and consequently classify an event (53-59). A superior tool in the pattern recognition process is the combination of a statistical and a structural approach (60, 61). As a result of this fact information provided by the WPR and WPC of a CAARMA model, which are statistical methods based on the LPR and LPC were combined with a structural approach to accurately identify the several waves associated with an event, and thus the event. So that the combined WPR/WPC and structural approach may best be understood, the logic behind a WLPR approach, resulting from a CAARMA model as well as the mathematical foundation of the WPR and WPC shall be established in this Section.

The basic form of the CAARMA model of TRAPS predicts the N^{th} windowed signal sample $x(n)$ by a linear combination of p excitation and q response time series. The differential outcome of the process, which in turn sets the foundation for predicting the next set of coefficients, is the prediction error e(n). Recalling the equations of Section 3, the

adaptive filter constructs a sequence of prediction coefficients a_k associated with the excitation and b_k associated with the response time series. The prediction residual is defined as the sum of the squares of the sequence e(n). The magnitude of this residual is both a function of the prediction coefficients $\{a(n), b(n)\}$, and the corresponding excitation and response time series. For a given data sequence $\{x_1(n)\}$ a set of prediction coefficients $\{a_1(n)\}$ can be calculated in association with $\{b_1(n)\}$ which minimizes the residual, e(n).

A method quite similar to this, but based on an AR method was discussed in Section 5 of the semi-annual report. It was quite successful, but the upgrade to the CAARMA model enhanced the method. There is a major conceptual difference between the prediction residual and coefficients of an AR and a CAARMA. The difference lies in the fact that the residual and coefficients of an AR model involve just $\{a(k)\}$, whereas the residual and coefficients of a CAARMA model involve both $\{a(k)\}$ and $\{b(k)\}$. The problem that arises from the AR to CAARMA transition is the relevance of $\{b(k)\}$ in the task of waveform classification. The relation of the CAARMA coefficients to the physical laws of the nuclear monitoring problem are discussed in Section 3.2. Once comprehension of the logic of Section 3.2 occurs, it becomes evident that the intent of $\{b(k)\}$ is to reproduce that

portion of the received signal which is the noise. Thus it is $\{a(k)\}$ that are uniquely related to an event.

Since the predictor coefficients $\{a_1(k)\}$ are uniquely related to the data sequence $\{x_1(n)\}$, they thus characterize that sequence. That is the coefficients maintain parity with the signature of the event. Now suppose another set of coefficients $\{a_2(k)\}$ are calculated by the CAARMA model on another data set $\{x_2(n)\}$. The degree of similitude between $\{a_2(k)\}$ and $\{a_1(k)\}$ determines whether or not the signatures of each event match. Unfortunately, much of the data is corrupted by artifacts and other irregularities. Thus despite the efficacy of the CAARMA model, certain coefficients and residuals can be corrupted. The resolution to this problem is the appropriate weighting of the residuals and coefficients. Hence the terminology WPR and WPC.

4.3.2 Segmentation for Structural Definition-Entropic Decision Process

The method of Section 4.3.1 established the mode by which the information provided by the CAARMA model can be statistically prepared for input for feed back into the range and bearing estimation technique. What remains to be disclosed is the mode used to elicit the structural information of the seismic wave. Seismic waveforms consist in segments of events of unknown duration, starting and ending points. Hence the technique, now presented, is based on the segmentation of the WPR.

The time trace provided by an array of geophones is a record of a number of events that occur after source energy is sent into the ground. The events include primary reflections (compressional and shear, multiples, surface waves, etc.). The first task is to isolate or segment these so that discrimination of surface waves against body waves, primaries against multiples, etc. can occur, since it is the information revealed by each of these events that is necessary to yield accurate range and bearing estimates. The WPR of Section 4.3.1 affords the segmentation we wish. Figures 31 and 32 give a visual illustration of how this segmentation occurs in a high SNR environment. Performance however in a low SNR is sensitive to the threshold selected. Some trial and error

was needed to obtain a proper threshold. Once segmentation and extraction of the segment transpired, the type of event now needed to be classified.

So that each type of event is accurately classified, the decision process which minimizes the probability of incorrect recognition must be determined. It is desirable to begin with a set of ideal measurements that extract only those parameters of the event which isolate distinguishing characteristics of the event type identity. Review of Section 3 equations and the discussion of Section 4.3.1 obviates the fact that the WPC physically represent the structure of the earth and consequently the distinguishing characteristics of the event type identity. A decision was thus made by determining both the simple average of the WPR of an event with a reference WPR and the entropic distance of the WPC.

A decision by a simple average of the WPR individual distances made a comparison between the reference and input contours results in a set of N individual distances d_1, d_2,..., d_n, then the overall distance D is given by

$$D = Yn \sum_{n=1}^{N} d_n$$

If D is less than some threshold distance B_t then the decision is made to accept the incoming WPR as a specific type of event. Otherwise, the WPR is rejected. Since the selection of contour

measurements and corresponding distances is based to a large extent on intuition the selection is not expected to be uniformly optimal. Therefore, a set of weights w_1, w_2,..., w_n, were used such that

$$D = \sum_{n=1}^{N} w_n d_n \qquad 4.13$$

where D best discriminates one event type from all the other event types.

As for the entropic distance, let a denote the p-dimensional column vector $(a_1,\ldots, a_p)$ and b denote the (p+1)-dimensional column vector $(1, a_1,\ldots, a_p)$. Now consider two adjacent sequences of measurement for reference and test sequences X_R and X_T of length N_R and N_T, respectively. The likelihood function under the null hypothesis that the two sequences are of the same event type is

$$L_o = (\hat{\sigma}_p \sqrt{2\pi})^{-(N_R' + N_T')} \exp[-1/2(N_R' + N_T')] \qquad 4.14$$

where

$$\hat{\sigma}_p^{\,2} = \hat{\underset{\sim}{a}}_p' \underset{\sim}{C}_p \hat{\underset{\sim}{a}}_p \qquad 4.15$$

and the poole l covariance matrix $\underset{\sim}{C}_p$ is given by

$$\underset{\sim}{C}_p = (N_R' \underset{\sim}{C}_R + N_T' \underset{\sim}{C}_T)/(N_R' + N_T'). \qquad 4.16$$

(62) The entropy distance d can then be defined as

$$d = -2 \ln (L_o/L_1) \qquad 4.17$$

which is a nonnegative number and symmetrical. If $\sigma_R = \sigma_T$ and $\hat{\vec{a}}_R = \hat{\vec{a}}_T$, then d = 0.

4.3.3 A Maximum Likelihood Feedback System

As cited in the semi-annual report, the original TRAPS employed a maximum likelihood (ML) method to solve the location problem. Particulars of the ML mathematics used can be found in Appendix B. The method of maximum likelihood (63) was introduced by the English mathematician, R.A. Fisher, as a general approach to the problem of estimating the "best" value for unknown parameters, $\vec{\Theta}$. If prior knowledge of the probability density of Θ is available, the likelihood function should be multiplied by this probability density before the maximization is carried out. It is assumed that the random parameter is restricted to a range of values and that the probability density of the parameter is constant in this range.

The arrival time, T_i, of the signal at the i^{th} geophone can be expressed by the following equation:

$$T_i = T + \frac{1}{v}\left[(x_i-x)^2 + (y_i-y)^2\right]^{\frac{1}{2}} + \varepsilon_i \qquad 4.18$$

where, x_i, y_i; are the (x,y) locations of the i^{th} geophone

x,y,T are the location and origin time of the event

V is the effective group speed of the signal

ε_i is an additive error terms assumed to be zero mean, Gaussian with known covariance matrix.

Determination of the arrival times at each geophone by the

method of Section 4.2, allows the estimators (x,y,T,V) to be chosen so as to maximize the likelihood.

$$L(\hat{x},\hat{t},\hat{T},\hat{V}) = \tfrac{1}{2} (\underset{\sim}{T} - \underset{\sim}{h}_i)^T \bigwedge_{\epsilon}^{-1} (T - h_i) \qquad 4.19$$

where

$$h_i \equiv T + \frac{1}{V} \left[(x_i - x)^2 + (y_i - y)^2\right]^{\frac{1}{2}} \qquad 4.20$$

However if $\hat{\underset{\sim}{A}} = (x, t, T, v)^T$ is one estimator of this sort, and $\hat{\underset{\sim}{B}}$ is another, then a composite minimum variance unbiased linear estimate $\hat{\underset{\sim}{C}}$ can be obtained as:

$$\bigwedge^{-1}_{\hat{\underset{\sim}{C}}} \hat{\underset{\sim}{C}} = \bigwedge^{-1}_{\underset{\sim}{A}} \hat{\underset{\sim}{A}} + \bigwedge^{-1}_{\hat{\underset{\sim}{B}}} \hat{\underset{\sim}{B}} \qquad 4.21$$

$$\bigwedge_{\hat{\underset{\sim}{C}}} = \bigwedge_{\hat{\underset{\sim}{A}}} + \bigwedge_{\hat{\underset{\sim}{B}}} \qquad 4.22$$

Thereby estimates derived from the independent wave components can be combined.

The investigation of this technique at the time of the semi annual report revealed this method was to have a gain significantly in excess of $\sqrt{N}$ for near field applications. Unfortunately, for far field cases where very weak events were detected much of the signal information pertaining to propagation

effects and origin was lost in the ambient noise; thus preventing Stage Three of TRAPS from achieving significant success. What was apparent was that the TRAPS process must be searched to understand the solution to this dilemma. Resolution came from two basic insights. The first insight was that the AARMA model contained information pertaining to the actual physics of the problem and effort had to be given to improve the acquisition of the information to insure extreme efficacy. Sections prior to this have indicated how this was accomplished. The second insight came from the realization that the closed loop system associated with the generation of the adaptive process is very structured. The existence of this underly ng feedback structure was extremely useful in developing an enhanced TRAPS' Stage Three. In essence what these insights allowed us to do was to first extract accurate information on the signal even in a low SNR environment and then use a feedback system to isolate the parameters that allow the ML method to be optimum, e.g. the processing system was in an iterative mode with the physics of the earth.

4.4 Technical Results

The results of this Section are based on the information made available by the CAARMA model, such as the WPR, WPC, etc. Thus discussion of these will be presented before the final range and bearing outcome of TRAPS.

The general mode of the WPR method, when used for event signature comparison is provided in Figures 31 and 32. Figure 32 is actually one of the best results obtained by the LPR method, but is provided since it graphically demonstrates the LPR and consequently WPR concept so extremely well. It is important to recall those parameters of the WPR, which have substantial effect on the results. The most obvious one is that of the hybrid adaptive model order and adaptation time used to generate the WPR. Despite the spectrum and consequently WPR variability generated by changing the order and adaptation time of the adaptive process, the variability of the WPR is highly consistent for each order group. Thus, as long as the same order is used to develop all the LPR classes for pattern recognition comparison, the probability of identification remains quite accurate. The second parameter is that of the number of residuals of which the WPR is summed. Whereas this factor is not critical in that there is a substantial range over which M can vary before results deteriorate, it is still an important issue. In general, we found

the summation factor of three times the order of the CAARMA process most effective. No substantial performance deterioration occurred between 1.0 and 50 times the order of the CAARMA process.

The coefficients of the CAARMA model are affected by the same factors as the WPR. Figures 33 and 34 are plots of the data that was processed by TRAPS to produce the coefficients found in Figures 35 and 36. Figures 37 and 38 are these same coefficients appropriately weighted for anomoly removal. These WPC and the associated WPR were then used to perform the necessarv range and bearing calculations. The range and bearing results can be found as follows:

Table 3 - Location results with original TRAPS model

4 - Location results with CAARMA model

5 - Location results with CAARMA model with improved time delay of Section 4.2

6 - Location results with same processing scheme as found in Table 5 with Maximum Likelihood feedback scheme added.

5.0 TASK 4: EXAMINE TRAPS FOR INCORPORATION INTO A COMPACT DIGITAL SYSTEM

When the signal processing research efforts progressed sufficiently to warrant the effort, consideration was given to design features for software and hardware requirements of a compact digital processing system. Since the realization of such a system within a real time limitation was possible with the original TRAPS (see Figure 1)(46), Task 4 efforts thus needed to concentrate solely on the additional software and hardware requirements imposed by the algorithm improvements detailed in this report. Due to the adaptive foundation of the final TRAPS with the CAARMA model (see Figure 3), such a compact system, that required no human intervention after initial reference waveform setting, was totally feasible. Unfortunately, without an array processor as part of the system, only near real time performance could be guaranteed.

6.0 IMPLICATIONS FOR FUTURE RESEARCH

A more robust TRAPS I for automatic detection, location and classification is possible. This could be accomplished first by means of research including Kalman formalism into TRAPS. This would allow the continuous upgrading of the model, providing time-varying spectral estimators that give superior results in a nonstationary environment. The goal of a more robust TRAPS I could also be met by incorporating into TRAPS a Spatial Dither Algorithm, which adjusts the signal to remove distortion. Since the signal is often obscured in a multi-path environment or by incomplete data, another future research project could address these problems by the development of a theory on the behavior of such parameters as partial correlation coefficients, pole/zero migration, etc., resulting from TRAPS I and the relationship of the behavior to the signature of the source of interest for improvement of the techniques presented in Section 4.3. The mathematical theory of the supposition proposed is based on the coupling of the theoretical concept of entropy from the fields of Information Theory and Thermodynamics. The culmination of the outcome of the above stated work could be produced onto a cognitive imaging system, such that TRAPS could be given the ability to "educate" itself on required input, thereby reducing human intervention.

REFERENCES

REFERENCES

1. B. Schnitta-Israel, D. Fletcher and J. Dorman, "Hybrid Adaptive Filtration for Seismic Event Location," Journ. Acoust. Soc. Am., Vol. 73, No. 1, pp230-241, Jan. 1983.

2. G.E.P. Box and G.M. Jenkins, Time Series Analysis: forecasting and control, Holden Day, San Francisco, pp 46-52, 1970.

3. E. Wegmen, Technometrics, Vol. 16, pp 321-322, 1974.

4. T.E. Landers and R.T. Lacoss, "Some Geophysical Applications of Autoregressive Spectral Estimates," IEEE Trans. Geos. Elect., Vol. GE-15, No. 1, pp 26-31, Jan. 1977.

5. T.J. Ulrych and T.N. Bishop, "Maximum Entropy Spectral Analysis and Autoregressive Decomposition," Reviews of Geophysics and Space Physics, pp 183-200, 1975.

6. H. Akaike, "A New Look at the Statistical Model Identification," IEEE Trans. on Automatic Control, Vol. AC-19, pp 716-723, December 1974.

7. T.W. Cairns, W.A. Cokerly, D.F. Findley, "ARMA Modeling Applied to Linear Prediction of Speech, IEEE CH1285, pp 23-26, June 1978.

8. B. Widrow and M.E. Hoff, "Adaptive Switching Circuits," IRE 1960 Wescon Conv. Rec., part 4, pp96-104.

9. L.J. Griffiths, "Rapid Measurement of Digital Instantaneous Frequency, IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-23, No. 2, pp 207-222, April 1975.

10. E.A. Robinson, Statistical Communication and Detection, New York: Hafner, 1967.

11. J.B. Thomas, An Introduction to Statistical Communications Theory, New York: John Wiley and Sons, 1969.

12. G.B. Whitham, Linear and Nonlinear Waves, New York: John Wiley and Sons, 1980.

13. E.A. Robinson and S. Treitel, Geophsical Signal Analysis, Englewood Cliffs, New Jersey: Prentice-Hall, 1980.

14. S.L. Marple, Jr., "Conventional Fourier, Autoregressive, and Special ARMA Methods of Spectrum Analysis," Engineer's Degree Thesis, Stanford University, 3 January 1977.

REFERENCES (cont.)

15. W.W. Shen, "A Constrained Minimum Power Adaptive Beamformer with Time-Varying Adaptation Rate," Geophysics, Vol. 44, No. 6, June 1979

16. B. Widrow, K. Duvall and T. Sake, 001AC Final Technical Report, Memistor Corporation, Stanford, California, ONR Contract No. N00014-82-C00420, March 28, 1983.

17. B. Widrow, "Adaptive Antenna Systems," Proc. of the IEEE, Vol. 55, No. 12, pp 2143-2159, December 1967.

18. A. Schuster, "On the Investigation of Hidden Periodicities with Application to a Supposed 26 Day Period of Meteorological Phenomena," Terrestrial Magnetism, Vol. 3, pp13-41, March 1898.

19. H.R. Radoski, P.F. Fougere and E.J. Zawalick, "A Comparison of Power Spectral Estimates and Applications of the Maximum Entropy Method," Journ. Geophysical Res., pp 619-625, Feb. 1975.

20. N. Anderson, "On the Calculation of Filter Coefficients for Maximum Entropy Spectral Analysis," Geophysics, Vol. 39, pp 69-72, Feb. 1974.

21. J.A. Cadzow, " Autoregressive Moving Average Spectral Estimation: A Model Equation Error Procedure," IEEE Trans. on Geos. Elect., Vol. GE19, No. 1, pp24-28, January 1981.

22. B. Schnitta-Israel, "Adaptive ARMA for Seismic Event Location," IEEE IGARSS, Wash. D.C., June 1981.

23. J.M Mendel, J. Kormylo, F. Aminzadeh, J.S. Lee and F. Habiki-Ashrafi, "A Novel Approach to Seismic Signal Processing and Modeling," Geophysics, Vol. 46, No. 10, pp 1398-1414, Oct. 1981.

24. C.L. Zahm, "Applications of Adaptive Arrays to Suppress Strong Jammers in the Presence of Weak Signals, Vol. AES-9, pp 260-271, 1973.

25. O.L. Frost, III, "An Algorithm for Linearly Constrained Adaptive Array Processing," Proc. of the IEEE, Vol. 60, No. 8, pp 926-935, August 1972.

26. W.F. Gebriel, "Adaptive Arrays - An Introduction," Proc. of the IEEE, Vol. 64, pp239-272, February 1976.

27. K. Duvall, "Signal Cancellation Phenomena in Adaptive Antennas: Causes and Cures," PhD dissertation, Stanford University, 1982.

28. B. Widrow, Et. al., "Adaptive Noise Cancelling: Principles and Applications," Proc. of the IEEE, Vol. 63, No. 12, pp 1692-1716, December 1975.

29. S.D. Stearns, "Error Surfaces of Recursive Adaptive Filters," IEEE Trans., CAS-28, Special Issue on Adaptive Signal Processing, June 1981.

30. L. Ljung, "Convergence of Recursive Estimators," Fifth IFAC Symposium on Identification and System Parameter Estimation, Darmstadt, 1979.

31. L. Ljung, "On Recursive Prediction Error Identification Algorithms," Report LiTH-ISY-I-0226, Dept of Electrical Engineering, Linkoping University, Sweden.

32. M.G. Larimore, J.R. Treichler and C.R. Johnson, Jr., "SHARF: An Algorithm for Adapting IIR Digital Filters," Vol. ASSP-28, No. 4, August 1980.

33. J.R. Treichler, M.G. Larimore and C.R. Johnson, Jr., "Simple Adaptive IIR Filtering," Proc. 1978 Intl. Conf. on Acoustics, Speech and Signal Proc., April 1978, pp 118-122.

34. M.G. Larimore, C.R. Johnson, Jr., and J.R. Treichler, "Adaptive Cancelling Using SHARF," Proc. 21st Midwest Symposium on Circuits and Systems, August 1978.

35. C.R. Johnson, Jr., M.G. Larimore and J.R. Treichler, "SHARF Convergence Properties," IEEE Trans. Acoustics, Speech and Signal Processing, Special Issue devoted to Adaptive Signal Processing, 1981.

36. J.R. Treichler, M.G. Larimore and C.R. Johnson, Jr., "On the Convergence Properties of the Simple Hyperstable Adaptive Recursive Filter (SHARF)," Proc. of the Intl. Conf. on Acoustics, Speech and signal Processing, April 1980.

37. D. Parikh, S.C. Sinha and N. Ahmed, "On a Modification of the SHARF Algorithm," Proc. of the 22nd Midwest Symposium on Circuits and Systems, June 1979.

38. A.V. Oppenheim and R.W. Schafer, Digital Signal Processing, Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1975.

39. J.P. Burg, "Three-Dimensional Filtering with an Array of Seismometers," Geophysics, Vol. 29, pp 693-713.

40. J. Capon, R.J. Greenfield and R.T. Lacoss, "Long-period Signal Processing Results for LASA," MIT, Lincoln Lab Technical Note, 1967-50.

41. F.H. Binder and J.P., "Wavenumber analysis of TFO Long-Noise Sample," Texas Instrument Special Report No. 17, 1966.

42. E.J. Douze, et al, "Short-period Seismic Noise," Bull. S.S.A., Vol. 57, pp 55-81.

43. R.A. Haubrich, "Semiannual Technical Summary- Study of Earth Noise on Land and Sea Bottem," University of California, Institute of Geophysics and Planetary Physics, February 1966.

44. D.R. Baumgardt, "Teleseismic P-Coda Stability and Coda Magnitude Yield Estimation," ENSCO, Semi-Annual Technical Report No. SAS-TR-83-01, August 1983.

45. S.F. Ingate, E.S. Husebye, A. Christoffersson, "Regional Arrays and Optimum Data Processing Schemes," NORSAR Contribution No. 348.

46. B. Schnitta-Israel, J. Sherry and Yosio Nakamura, "Hostille Artillery Location and Classification Using a Hybrid Adaptive Processing System," ONR Final Report, Contract N00014-77-C-0446, 1985.

47. IEEE Trans. Acoustical, Speech and Signal Processing, Special Issue on Time Delay Estimation, June 1981.

48. B. Schnitta-Israel and J. Sherry, "Nuclear Monitoring, Using a Hybrid Adaptive Processing System, TRAPS I," DARPA Contract No. F49620-83-C-0137 Semi annual report, February 14, 1984.

49. Y.T. Chan, J.M. Riley and J.B. Plant, "A Parameter Estimation Approach to Time Delay Estimation and Signal Detection," IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-28, February 1980.

50. Y.T. Chan and R.K. Miskowicz, "Estimation of Coherence and Time Delay with ARMA Models," IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-32, No. 2, April 1984, pp295-303.

51. M. Morf, et al, "Investigation of New Algorithms for Locating and Identifying Spatially Distributed Sources and Receivers," Tech. Summary Rep. to DARPA, Report M355-1, March 1979.

52. B. Porat and B. Friedlander, "Estimation of Spatial and Spectral Parameters of Multiple Sources," IEEE Trans. Inform. Theory, Vol. IT-29, pp 412-425, May 1983.

53. R.V. Churchill, Complex Variables and Applications, New York: McGraw-Hill, 1960.

54. B.S. Atal, "Effectiveness of Linear Prediction Characteristics of the Speech Wave for Automatic Speaker Identification and Verification," J. Acoust. Soc. Am., Vol. 55, NO. 6, pp 1304-1312, June 1974.

55. H. Harjes, "Spektral Analytische Interpretation Seismischer Aufzeichnungen," Geologisches Jahrbuch, Reihe E. Heft 17, Hannover, 1979.

56. F. Itakura, "Minimum Prediction Residual Principle Applied to Speech Recognition," IEEE Trans. Acoust., Speech and Signal Processing, ASSP-23, pp 67-72, February 1975.

57. P. Bois, "Autoregressive Pattern Recognition Applied to the Delimitation of Oil and Gas Reservoirs," Geophysical Pros., Vol. 28, pp 572-591, 1980.

58. A.E. Rosenberg and M.R. Sambur, "New Techniques for Automatic Speaker Verification," IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-24, pp 169-176, April 1975.

59. O. Sandrin and D. Tjostheim, "Multivariate Autoregressive Representation of Seismic P-Wave Signals with Application to Short-Period Discrimination," Bull. S.S.A., Vol. 68, No. 3, pp 735-756, June 1978.

60. C.H. Chen, "Geophysical Signal Recognition," Technical Report SMU-EE-TR-81-5, So. Mass. Univ., January 29, 1981.

61. B. Schnitta-Israel, "Hybrid Image Processing System by Means of an ARMA Algorithm," IEEE Trans, Vol. SU-31, No.4, pp 391-395, July 1984.

62. C.H. Chen, "Pattern Analysis of Acoustical and Seismic Events," Proc. Intl. Symp. Computer Aided Seismic Analysis and Discrimination, IEEE 83CH1890-3, pp 114-117, June 1983.

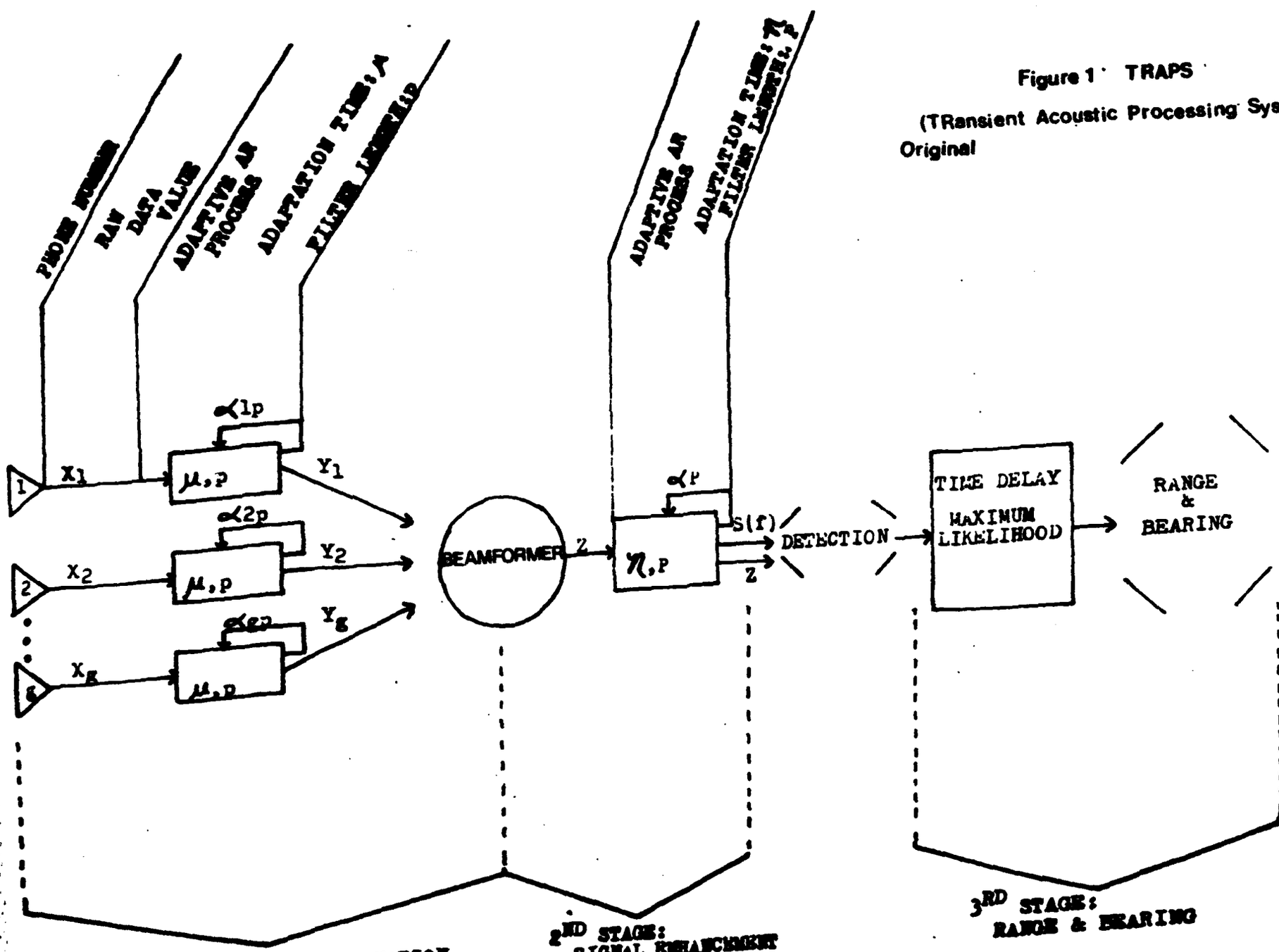
63. H. Cramer, "On the Theory of Stationary Random Process, Ann. Math, Vol. 41, pp 215-230, 1940.

FIGURES

Figure 1 TRAPS

(TRansient Acoustic Processing System)

Original



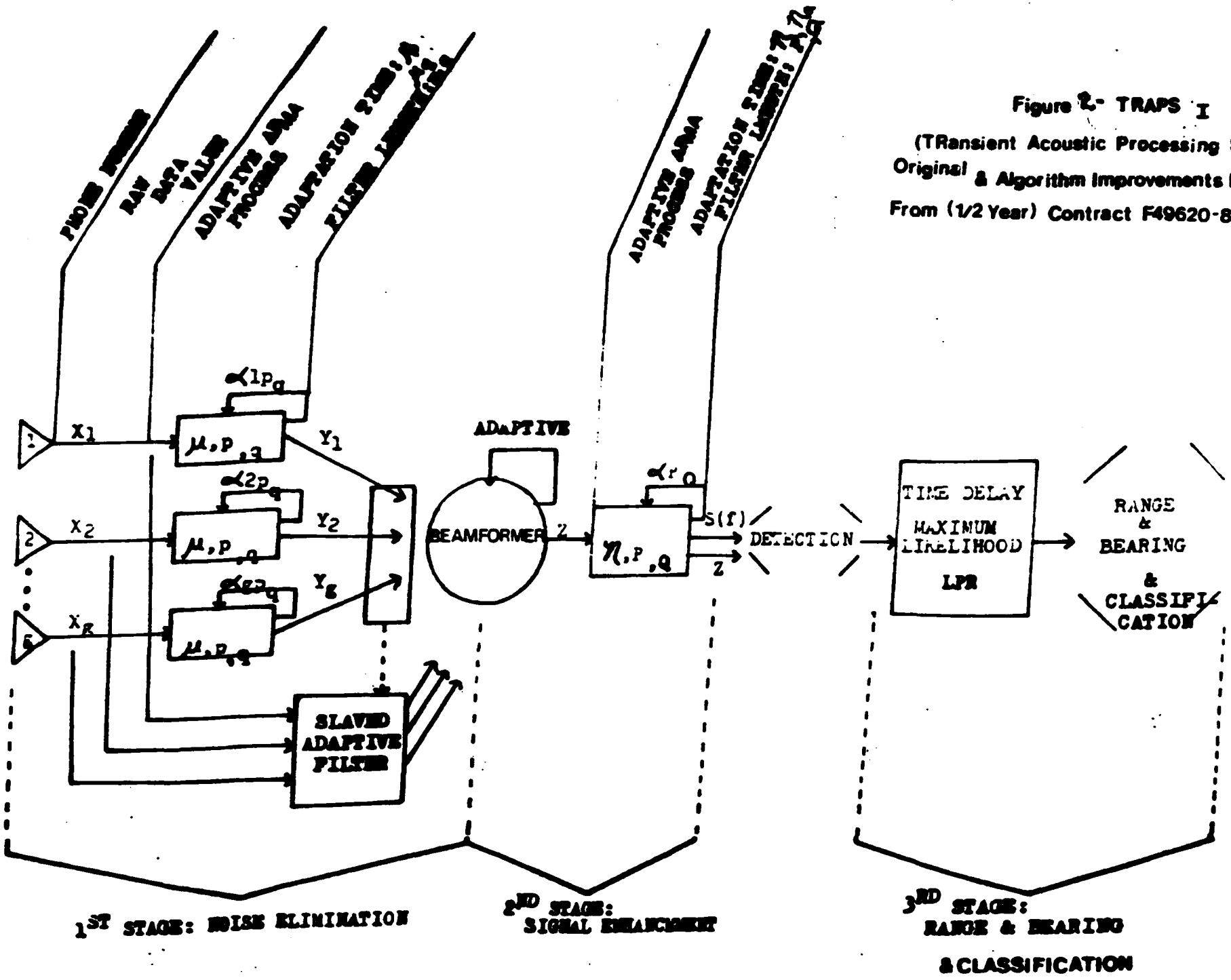


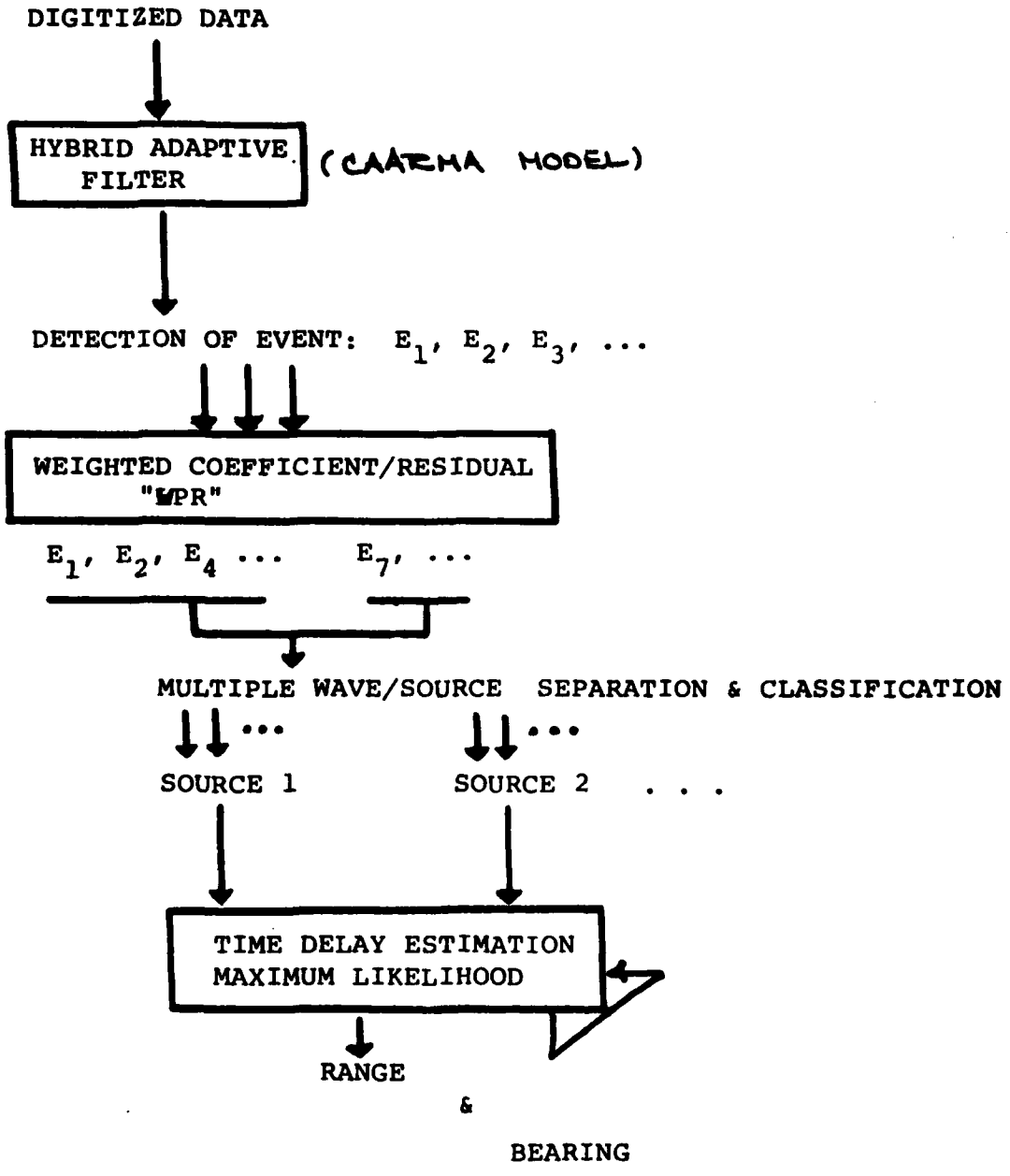
Figure 2- TRAPS I

(TRansient Acoustic Processing System)

Original & Algorithm Improvements Resulting

From (1/2 Year) Contract F49620-83-C-0137

Figure 3. TRAPS Schematic



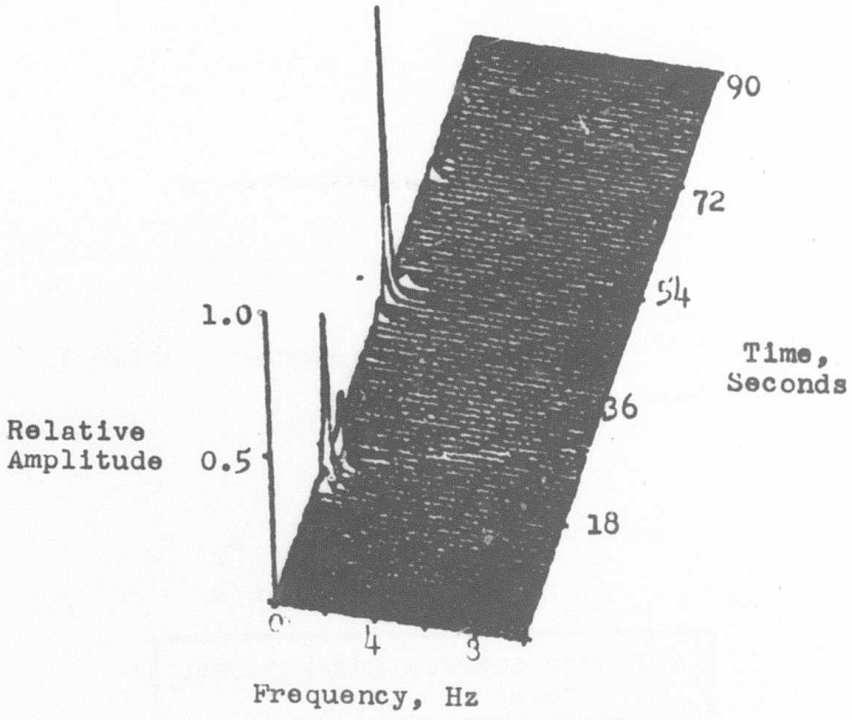


Figure 4. Instantaneous Power Spectrum, resulting from TRAPS' Stage Two only, using a filter length of 4. This plot was the product of a simulation study specifically designed for a TRAPS' Stage One processing gain demonstration. The simulation study consisted of 3 seismic waves each with a frequency transition from 5.2 Hz to 4.5 Hz occurring over the 8 second length of the signal. Further conditions placed on each wave were as follows:

Signal 1: 5 db
begins at second 17
Signal 2: 10 db
begins at second 42
Signal 3: 5 db
begins at second 71

Superimposed on these signals were three .5 Hz, 10 second length seismic waves, we shall term noise, coming from the same direction as the signals. Further conditions placed on each wave were as follows:

Noise 1: 15 db
begins at second 17
Noise 2: 40 db
begins at second 42
Noise 3: 10 db
begins at second 71

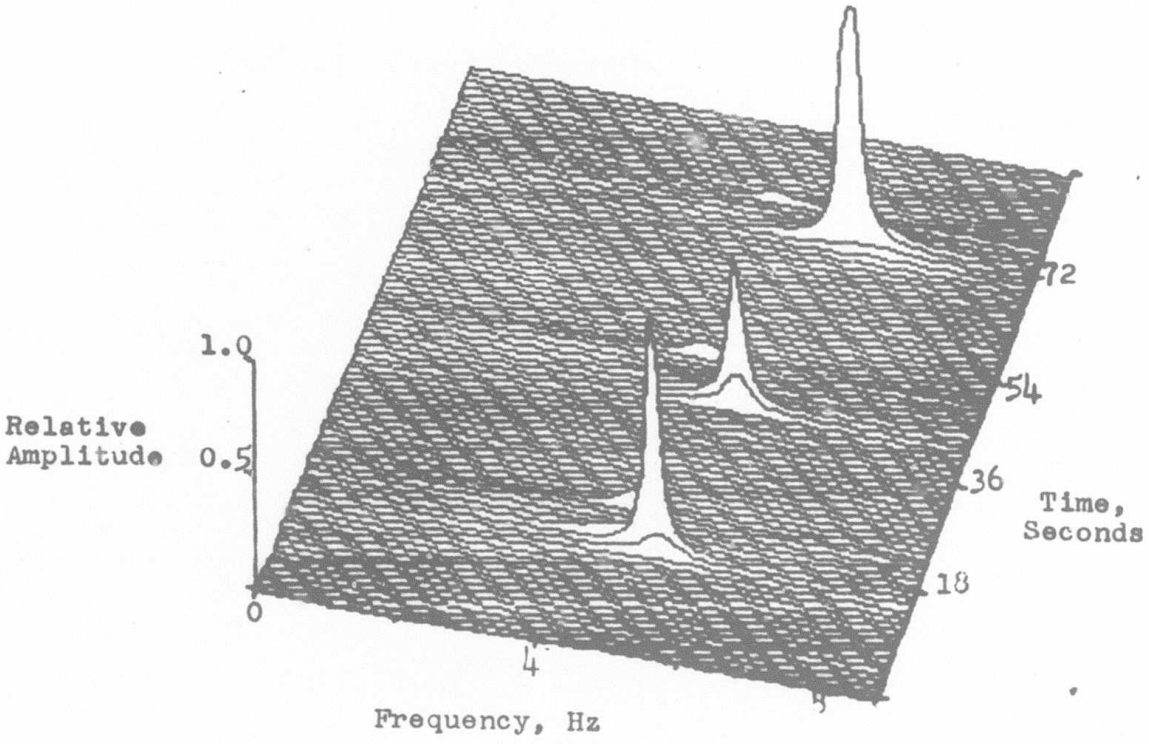


Figure 5. Instantaneous Power Spectrum, resulting from TRAPS' Stage One and Stage Two, using a filter length of 4 for each stage. This plot was the product of a simulation study specifically designed for a TRAPS' Stage One processing gain demonstration. The simulation is that as detailed on Figure 4 caption.

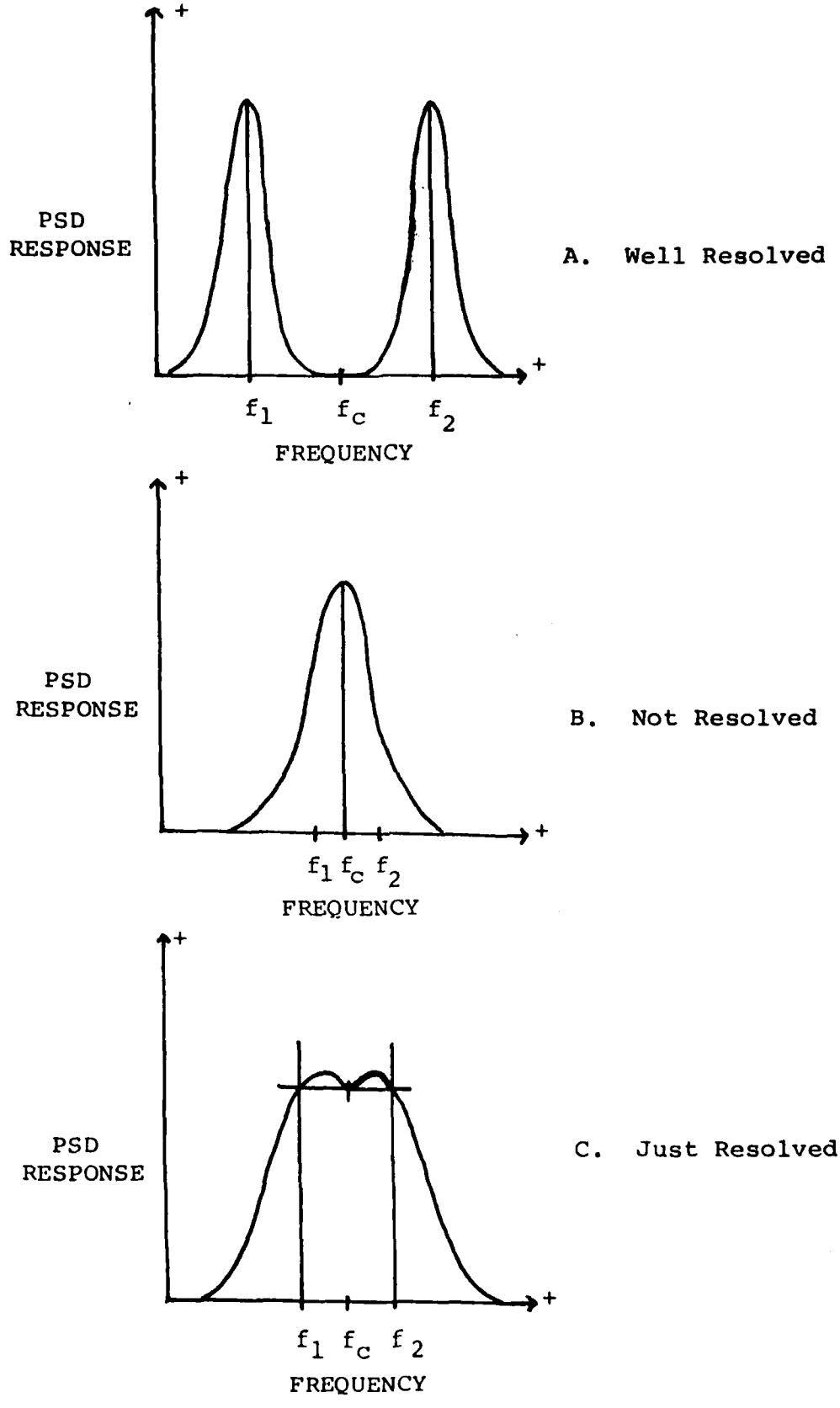


Figure 6. Condition of Power Spectral Density (PSD) Resolution for Frequency Resolution Study; the results of which are given in figures 7 and 27.

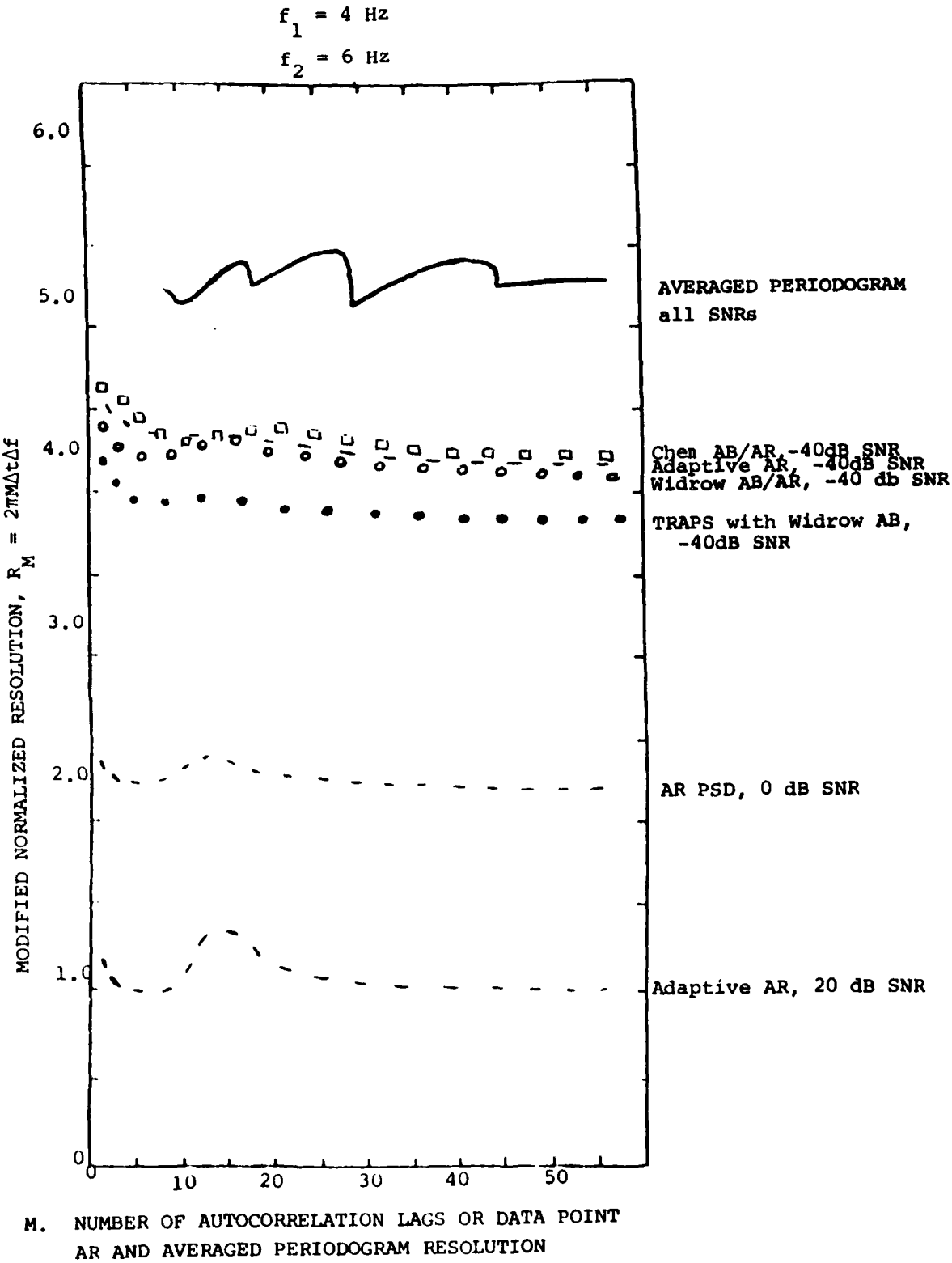


Figure 7. Results of Adaptive Beamformer (AB) Frequency Resolution Evaluation, Note that a smaller R denotes improved resolution.

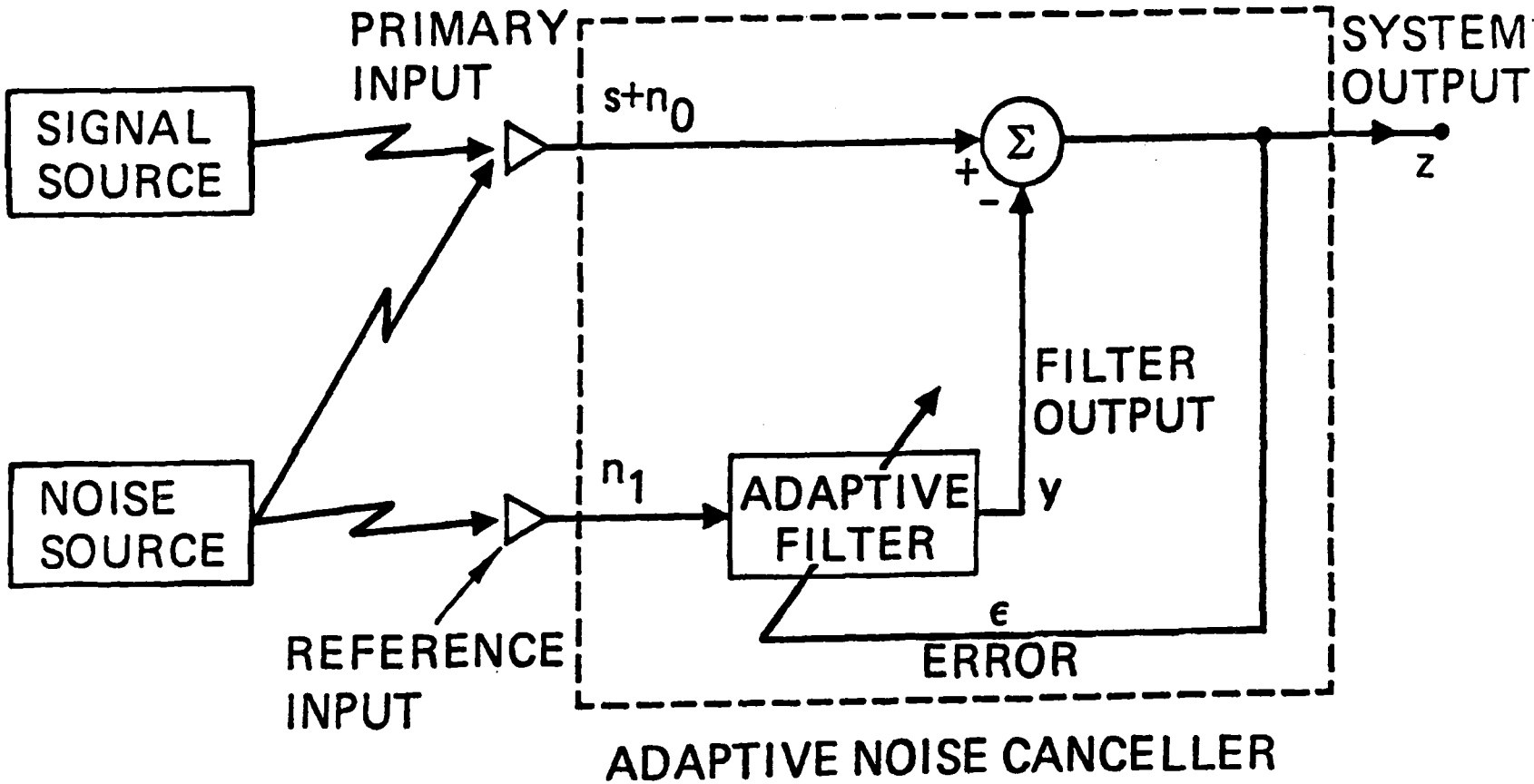
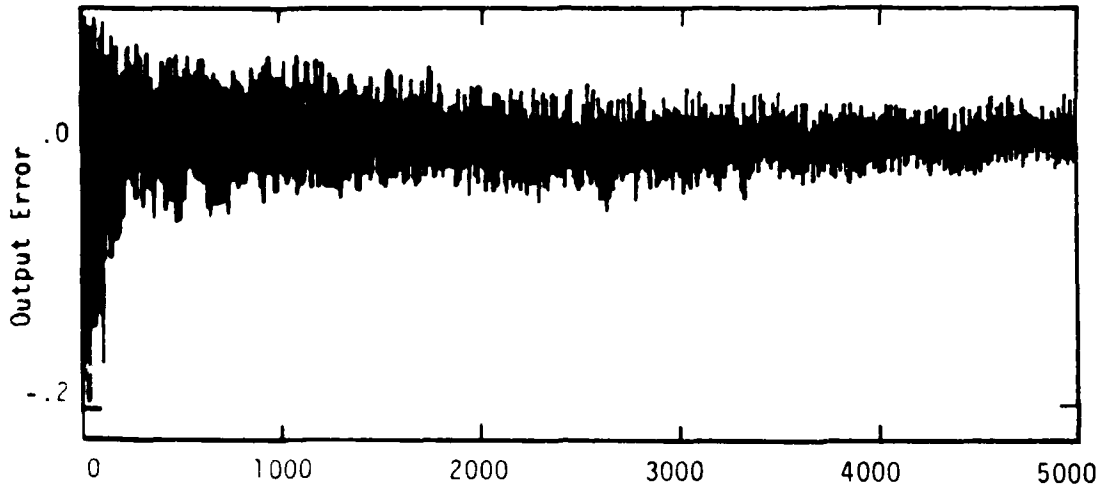
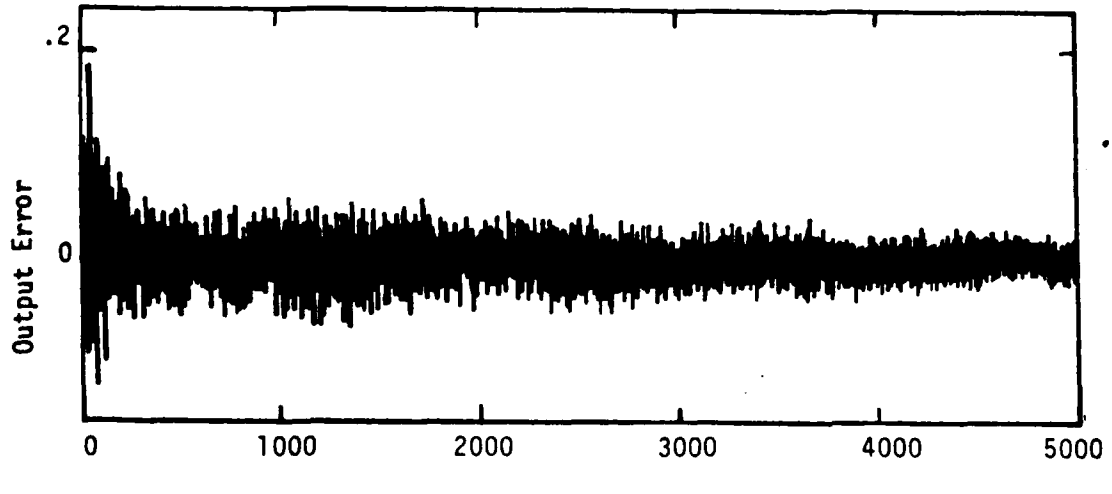


Figure 8. TRAPS 1st Stage noise cancelling concept.



a) $\hat{a}_{12}(0) = -1.5$; $\hat{a}_{22}(0) = -.8$



b) $\hat{a}_{12}(0) = 1.5$; $\hat{a}_{22}(0) = -.8$

Figure 23. Insufficient Order Adaptive Algorithm.

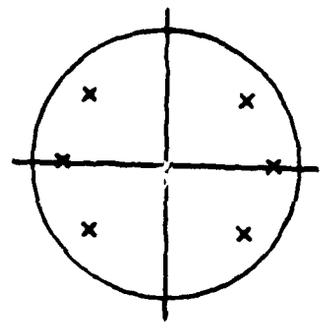
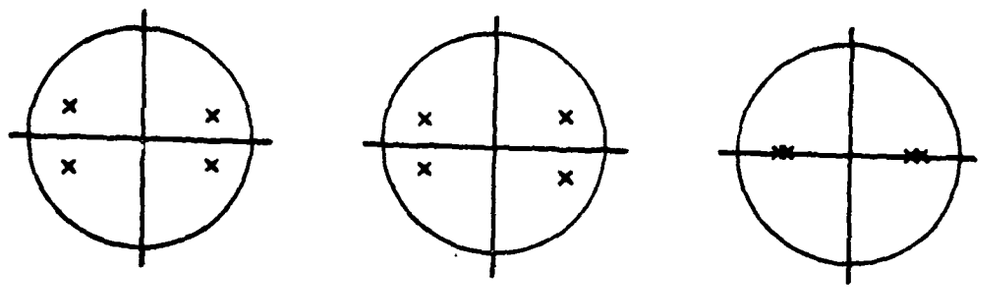


Figure 21. Model Pole Locations in Z-Plane.

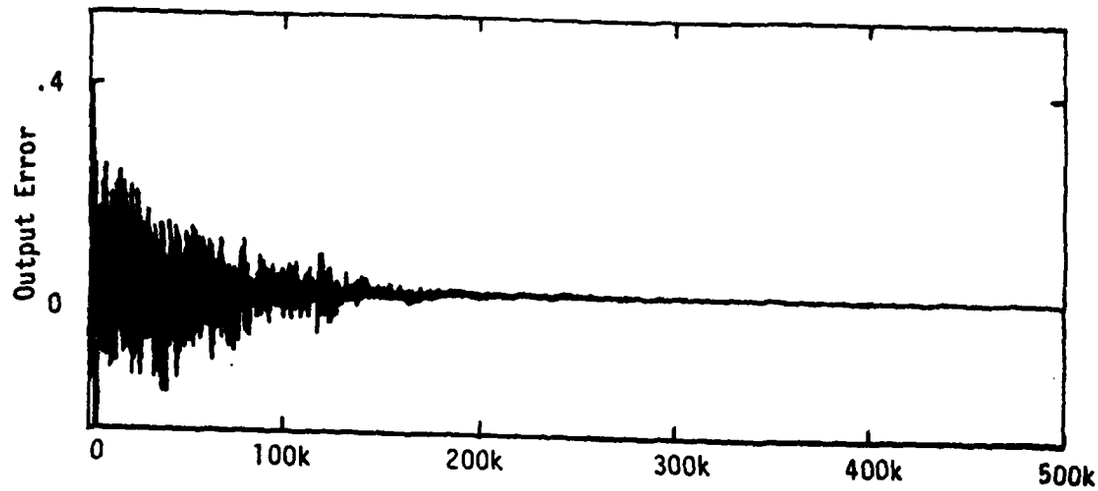


a) $\hat{a}_{12}(0) = -1.5$
$\hat{a}_{22}(0) = -.8$

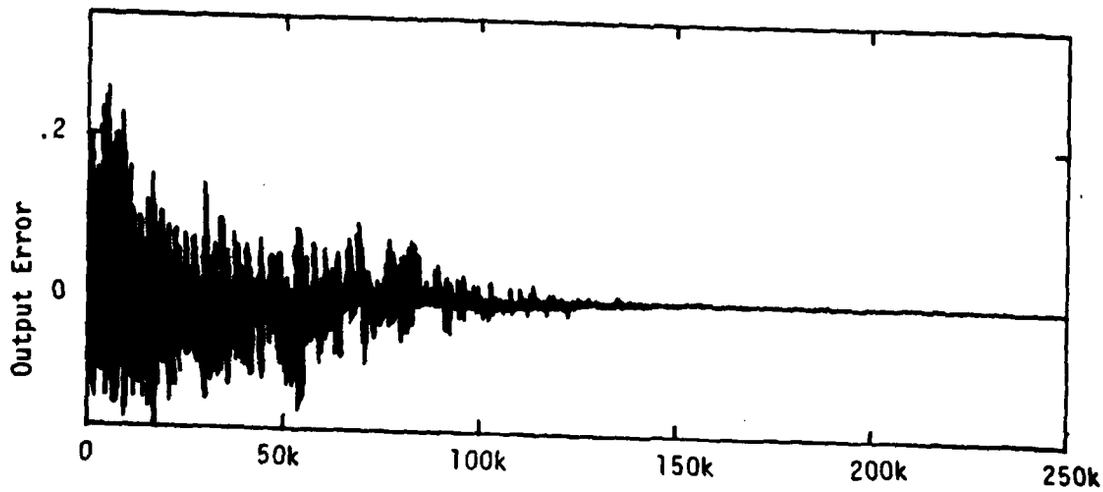
b) $\hat{a}_{12}(0) = 1.5$
$\hat{a}_{22}(0) = -.8$

c) $\hat{a}_{12}(0) = .0$
$\hat{a}_{22}(0) = .8$

Figure 22. Adaptive Algorithm Pole Locations.

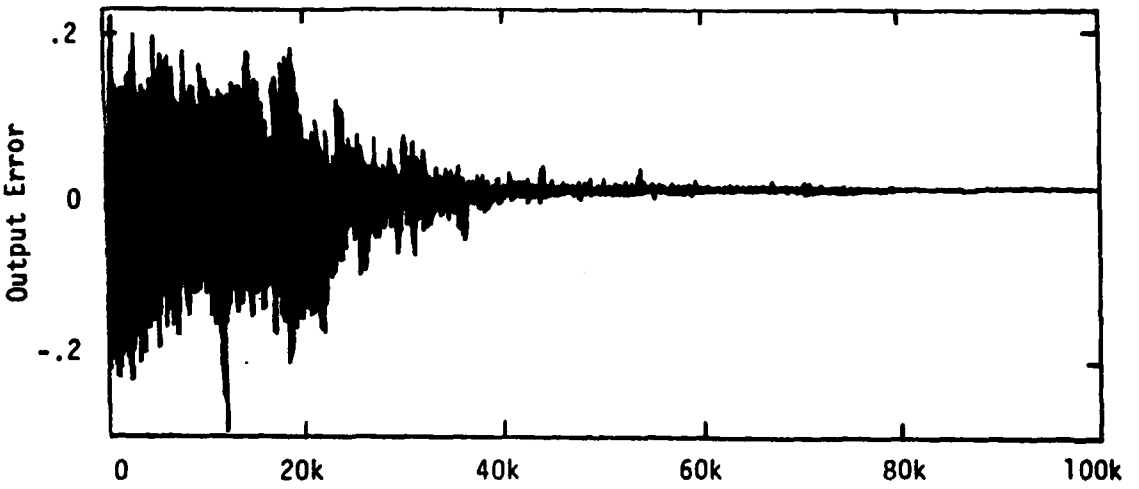


a) RLMS Algorithm

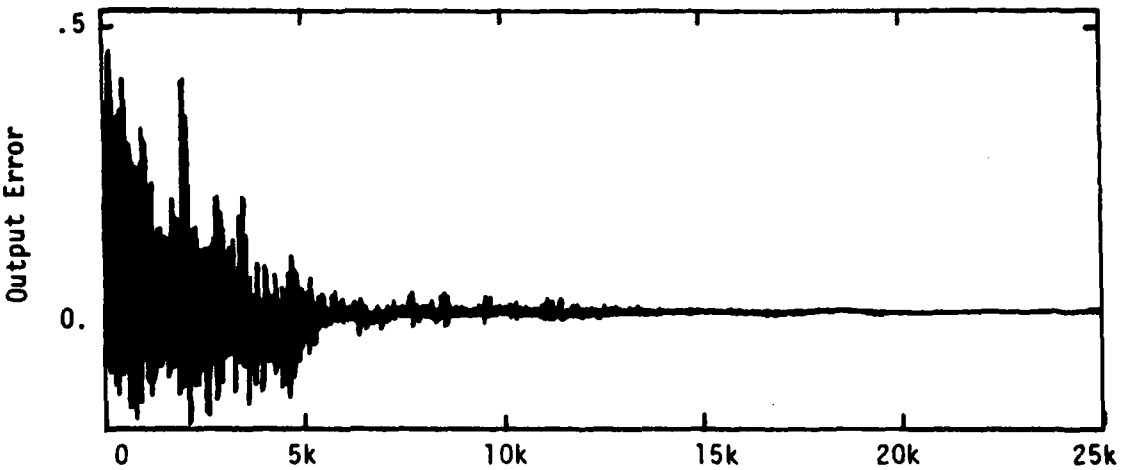


b) SHARF Algorithm

Figure 20. Fourth Order Example: Direct Form Algorithms.

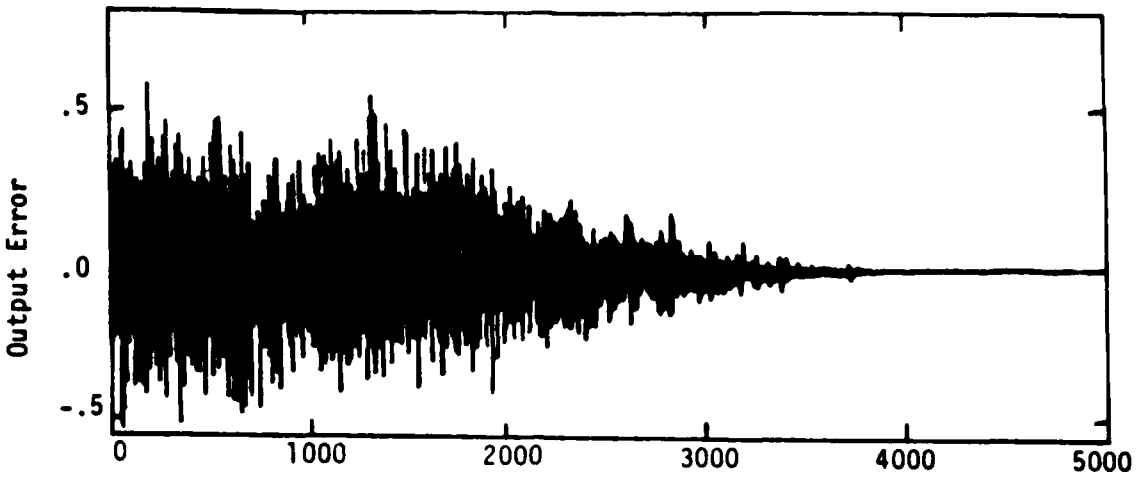


a) CRLMS Algorithm

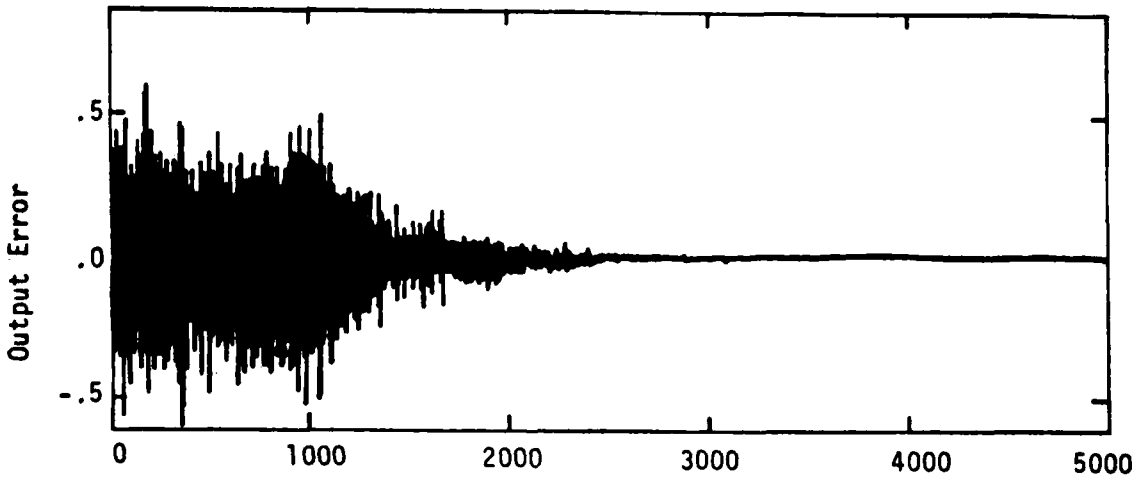


b) NCRLMS Algorithm

Figure 19. Fourth Order Example: Cascade Structure Algorithms.

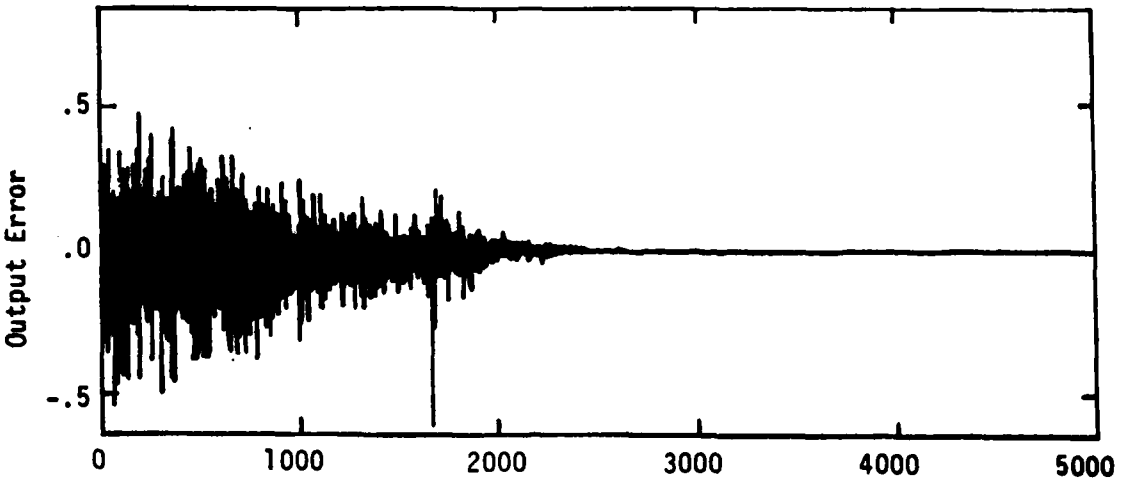


a) $C(z) = 1$

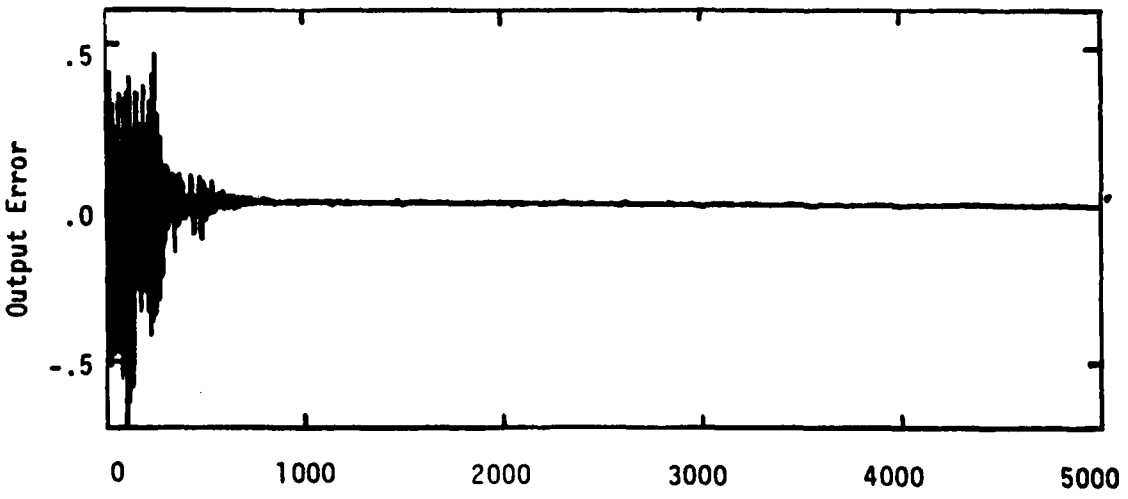


b) $C(z) = 1. - 1.25z^{-1}$

Figure 18. Second Order Example: SHARF Algorithm.

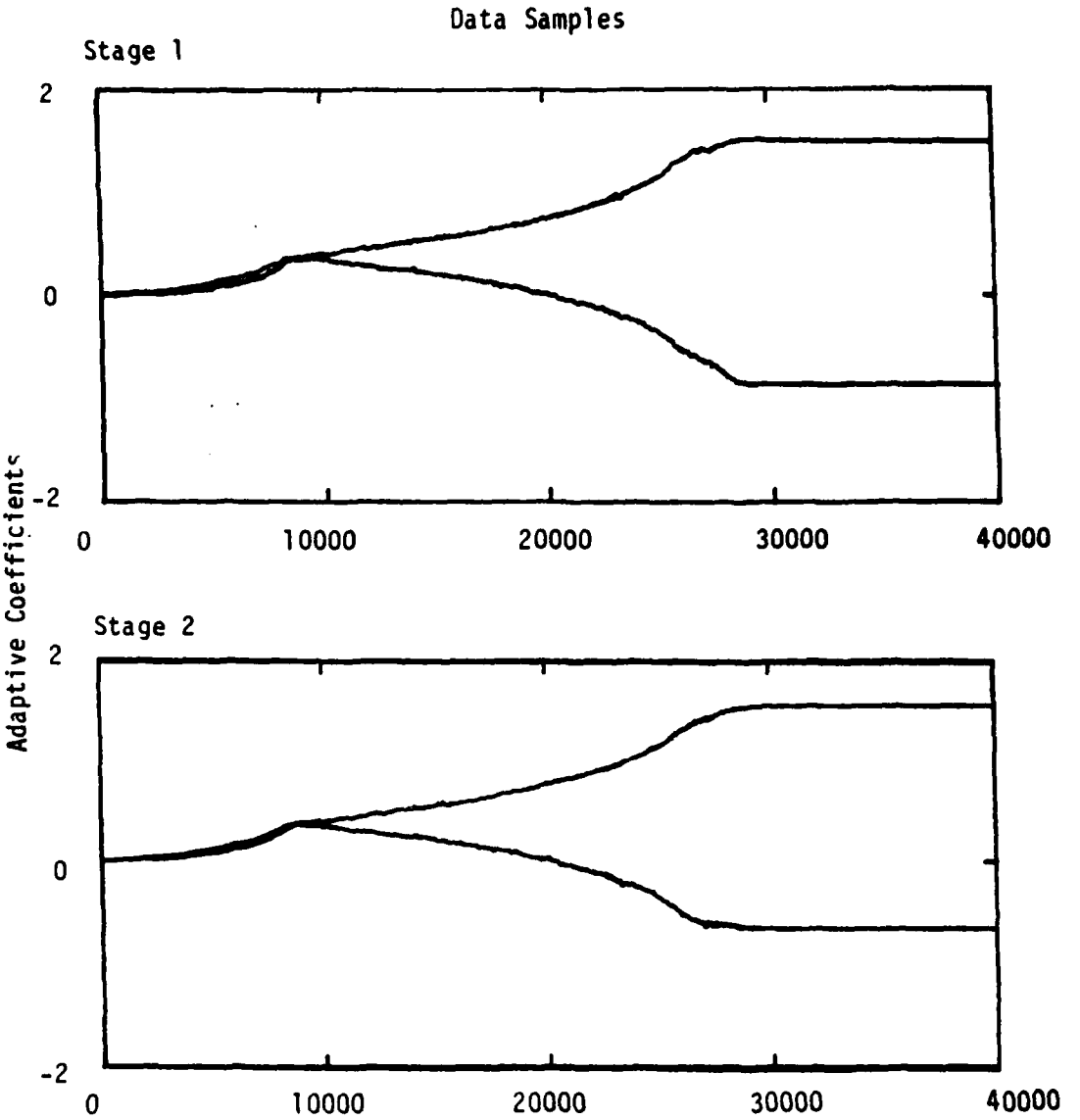


a) RLMS Algorithm



b) CAARMA Algorithm

Figure 17. Second Order Example: Gradient Search Algorithms.



Model:

$a_{11} = 1.5550 \quad a_{21} = -.6490$

$a_{12} = 1.4990 \quad a_{22} = -.8480$

Adaptive Parameters:

$\hat{a}_{11} = 1.4990 \quad \hat{a}_{21} = -.8480$

$\hat{a}_{12} = 1.5549 \quad \hat{a}_{22} = -.6489$

Figure 16. Fourth Order CRLMS Example.

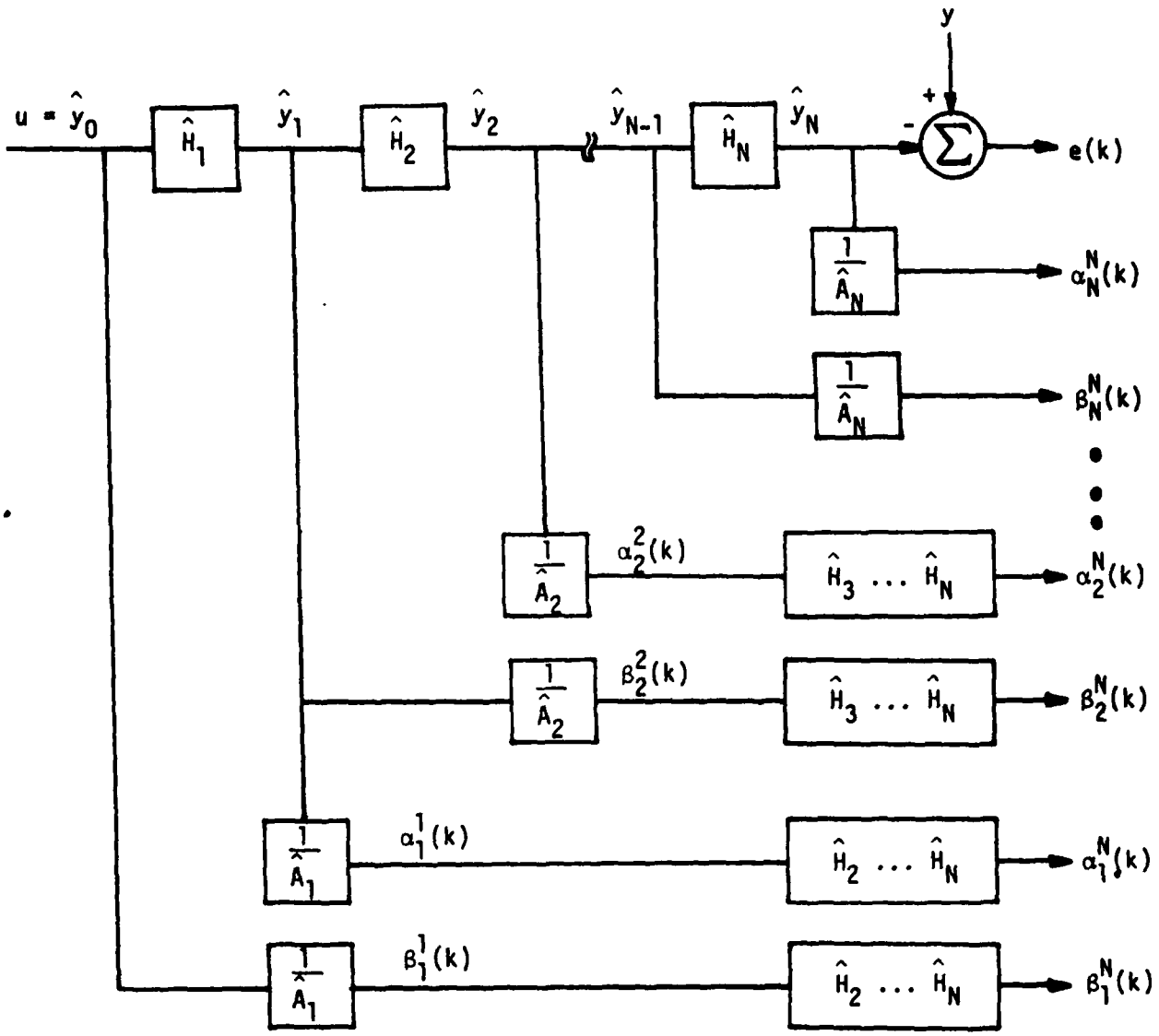


Figure 15. CAARMA Gradient Formulation.

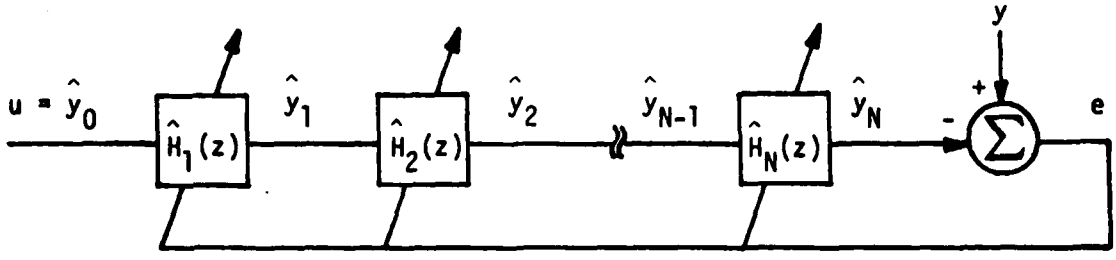
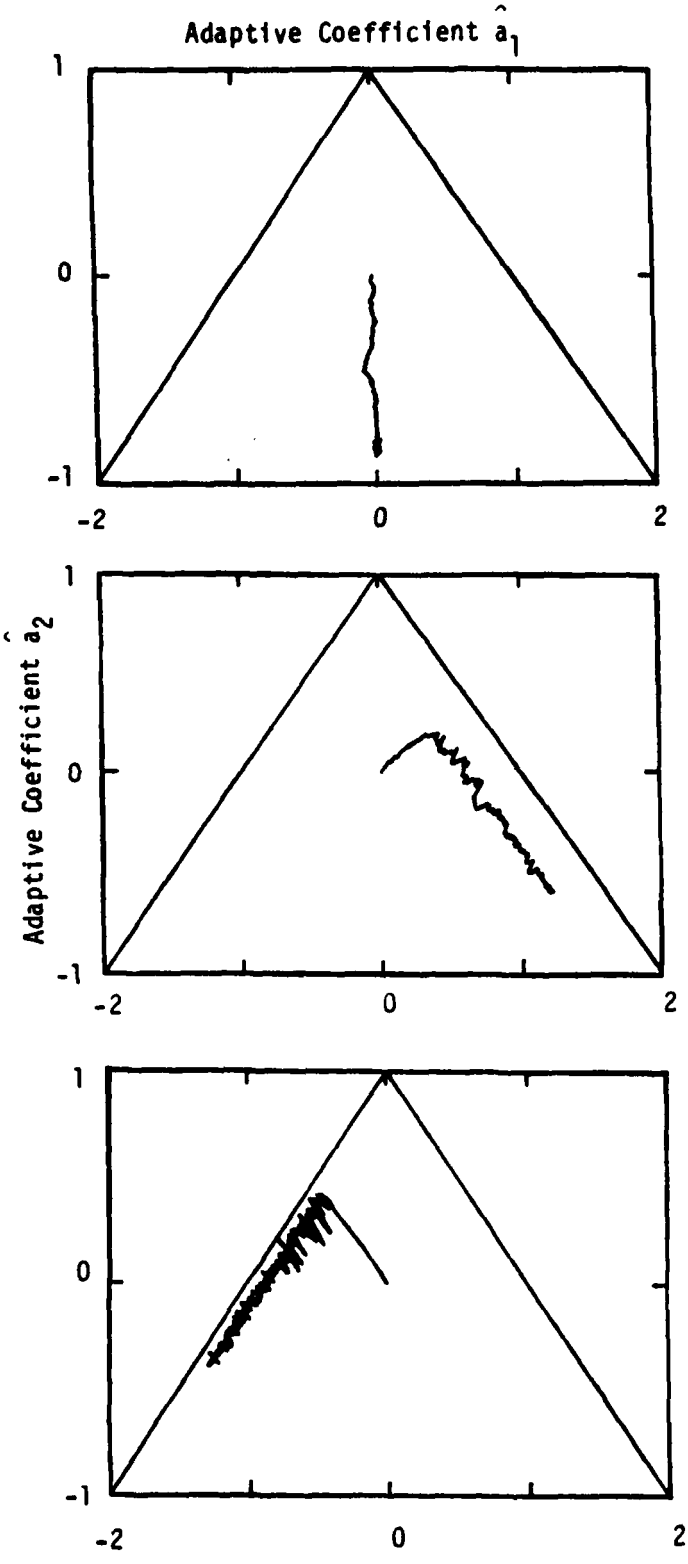


Figure 14. Adaptive Cascade Structure.



a) Model:

$a_1 = .0000 \quad a_2 = -.8000$

Adaptive Parameters:

$\hat{a}_1 = -.0021 \quad \hat{a}_2 = -.8010$

(After 300 samples)

$\rho = .01$

b) Model:

$a_1 = 1.2000 \quad a_2 = -.6000$

Adaptive Parameters:

$\hat{a}_1 = 1.1997 \quad \hat{a}_2 = -.5998$

(After 800 samples)

$\rho = .01$

c) Model:

$a_1 = -1.3000 \quad \hat{a}_2 = -.4000$

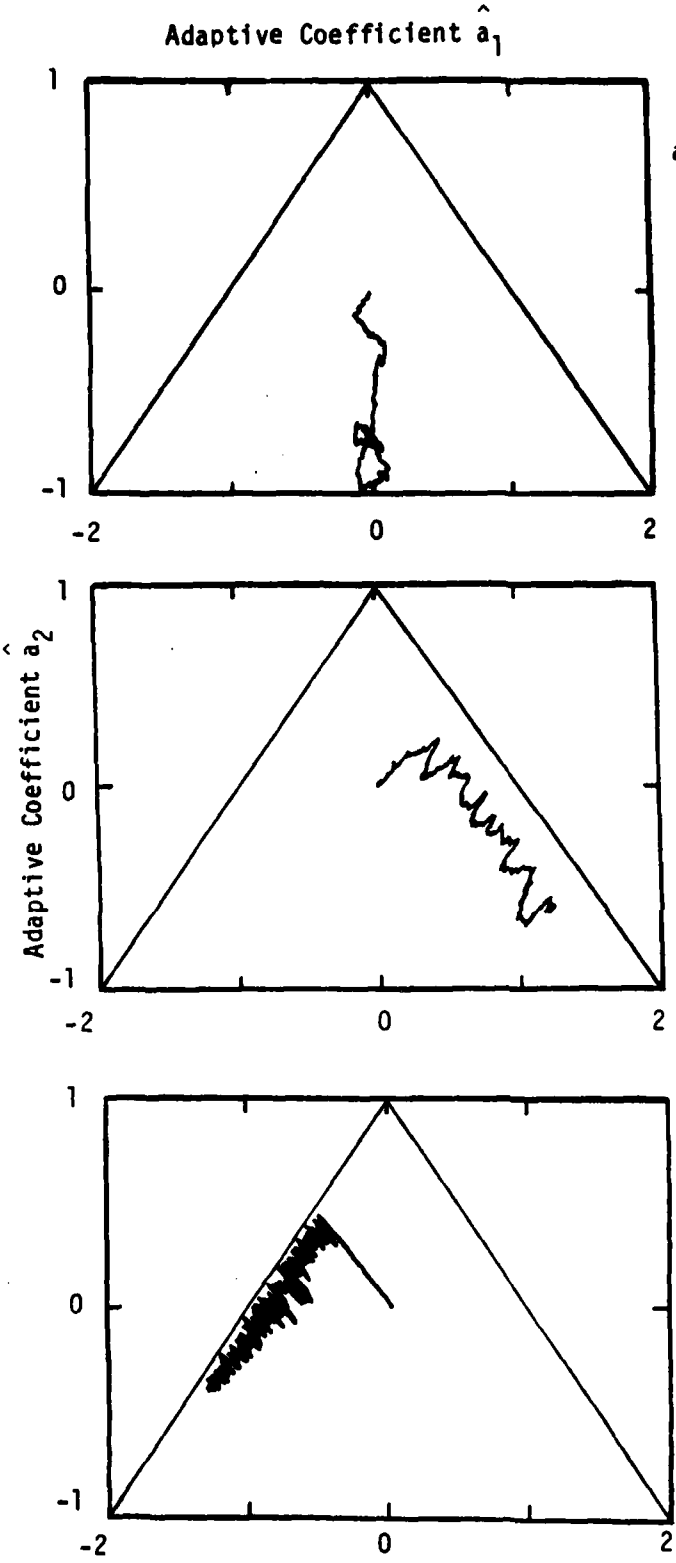
Adaptive Parameters:

$\hat{a}_1 = 1.2995 \quad \hat{a}_2 = -.3996$

(After 10,000 samples)

$\rho = .0005$

Figure 13. RLMS Parameter Tracks.



a) Model:
$a_1 = .0000 \quad a_2 = -.8000$

Adaptive Parameters:
$\hat{a}_1 = -.0012 \quad \hat{a}_2 = -.8002$
(After 120 samples)
$\rho = .05$

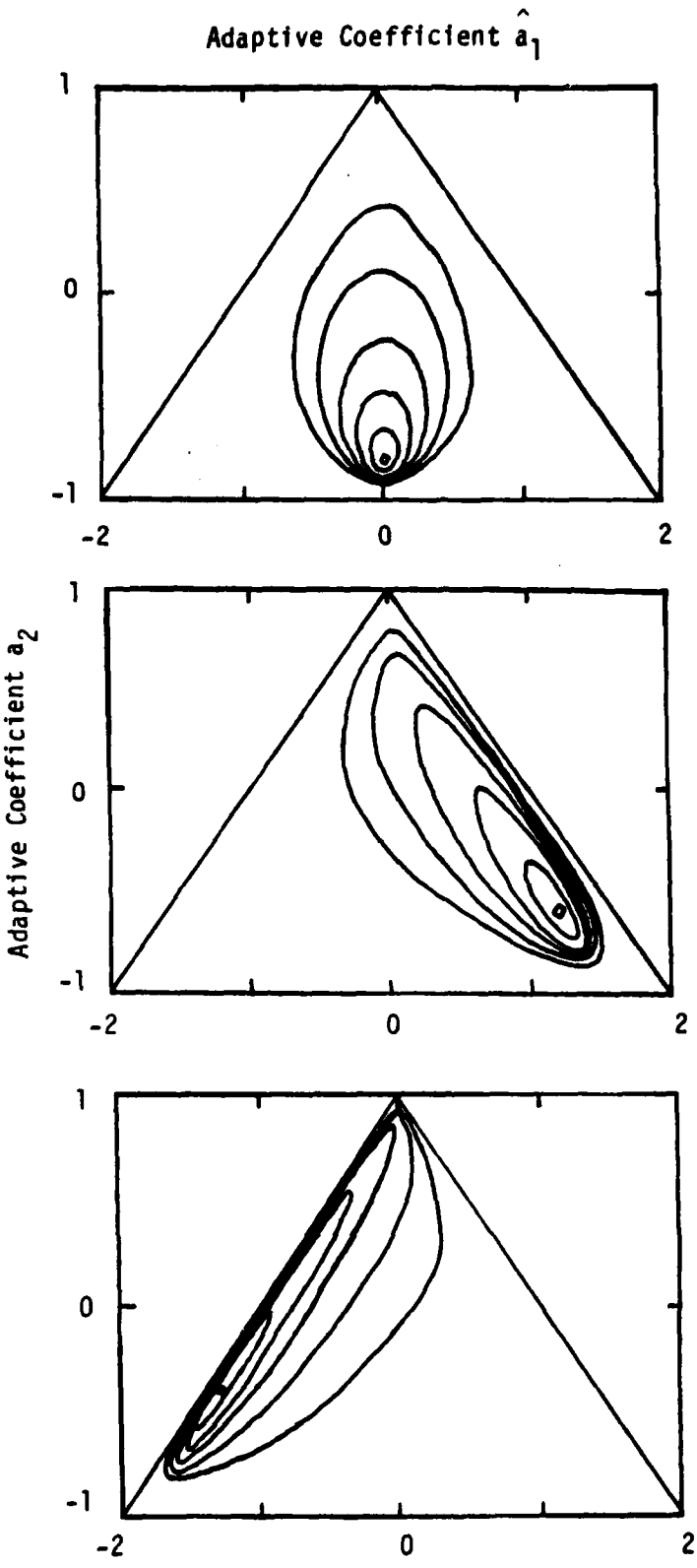
b) Model:
$a_1 = 1.2000 \quad a_2 = -.6000$

Adaptive Parameters:
$\hat{a}_1 = 1.2000 \quad \hat{a}_2 = -.5997$
(After 400 samples)
$\rho = .025$

c) Model:
$a_1 = -1.3000 \quad a_2 = -.4000$

Adaptive Parameters:
$\hat{a}_1 = -1.2996 \quad \hat{a}_2 = -.3995$
(After 5600 samples)
$\rho = .001$

Figure 12. RLMS Parameter Tracks.



a) Imaginary Poles

$a_1 = 0.0$

$a_2 = -.8$

b) Complex Poles

$a_1 = 1.2$

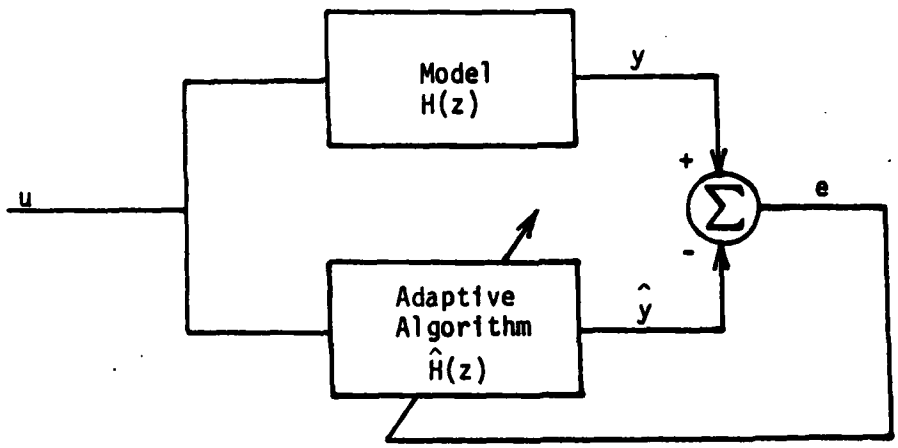
$a_2 = -0.6$

c) Real Poles

$a_1 = -1.3$

$a_2 = -0.4$

Figure 11. Performance Surface Examples.



System Identification Configuration.

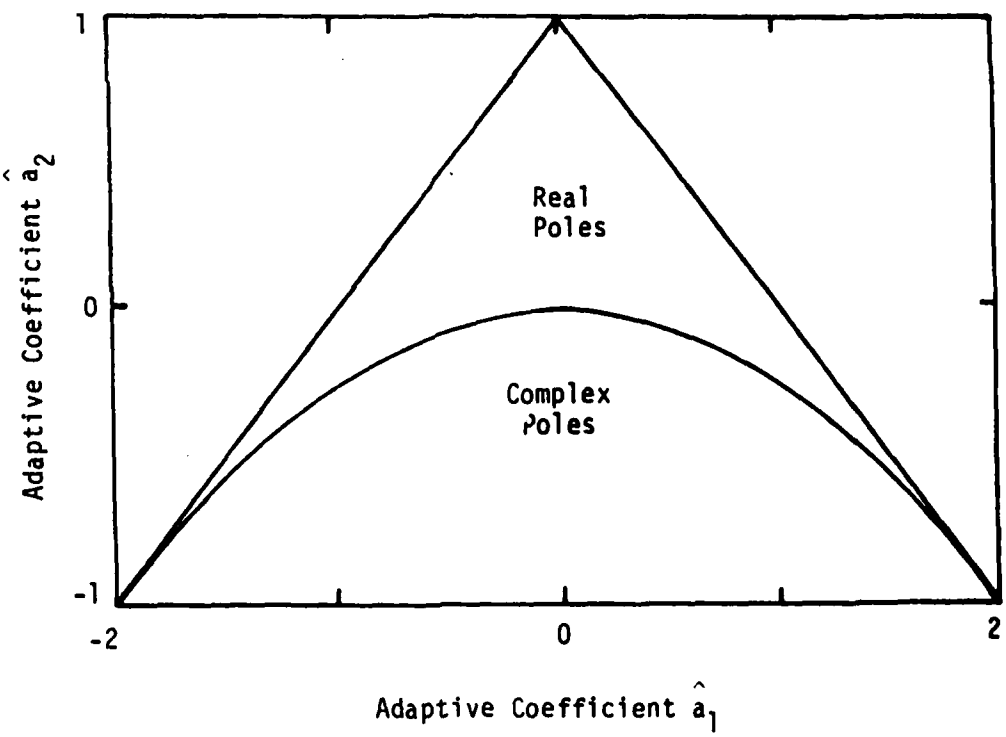


Figure 10. Parameter Stability Triangle.

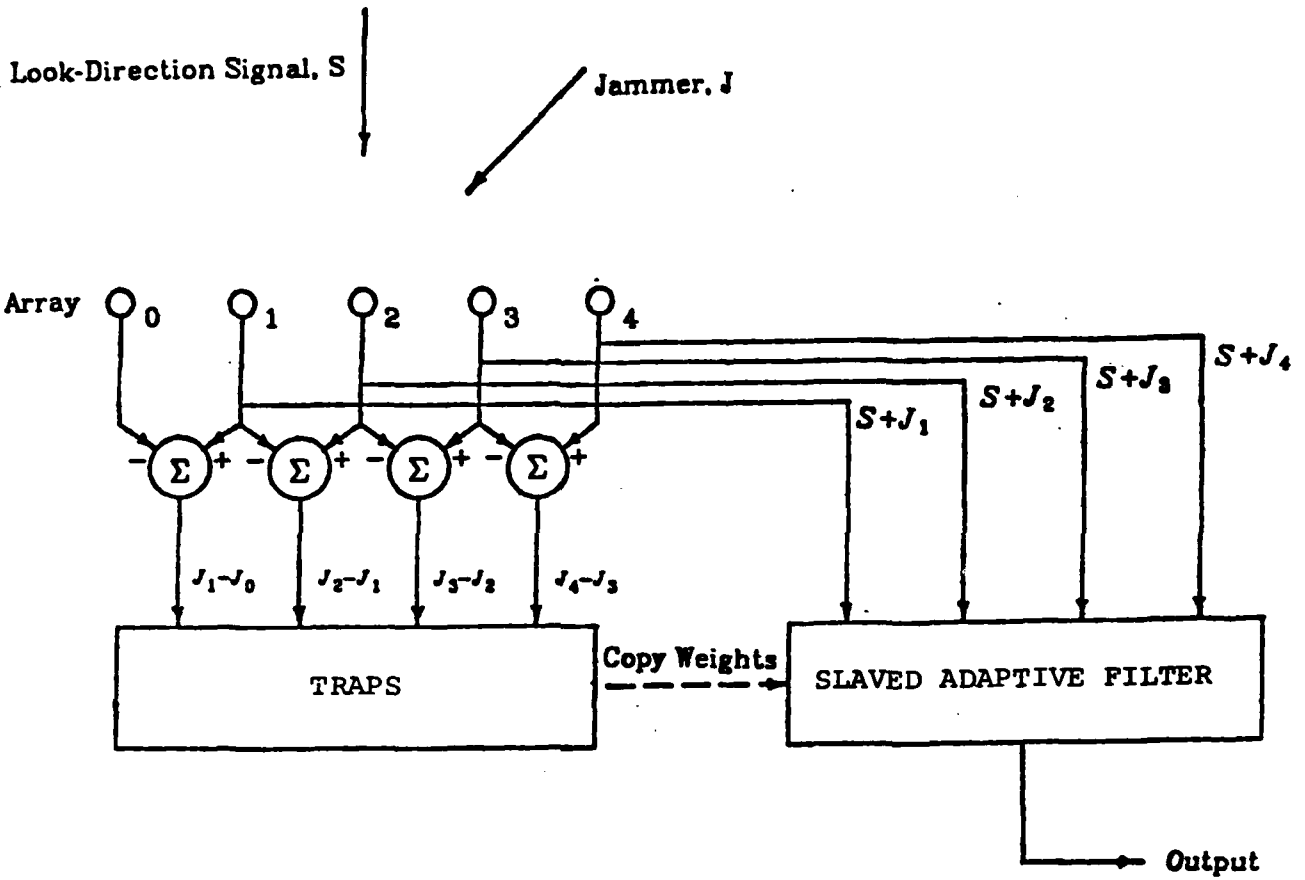
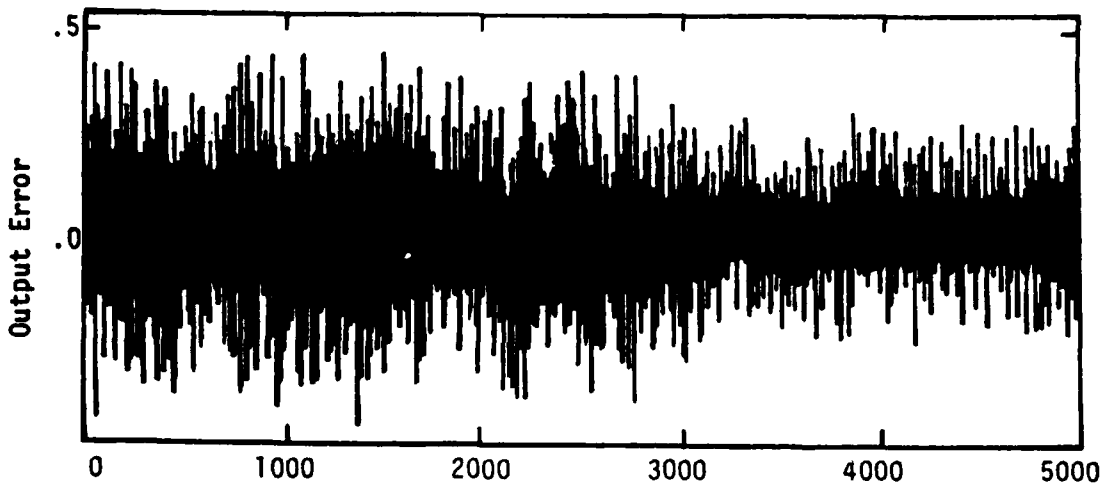


Figure 9. Duvall system for preventing signal cancellation phenomena



c) $\hat{a}_{12}(0) = .0$; $\hat{a}_{22}(0) = .8$

Figure 23. Insufficient Order Adaptive Algorithm.

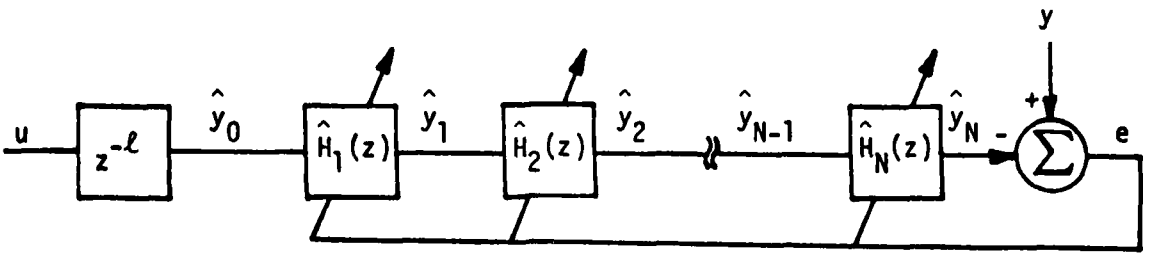
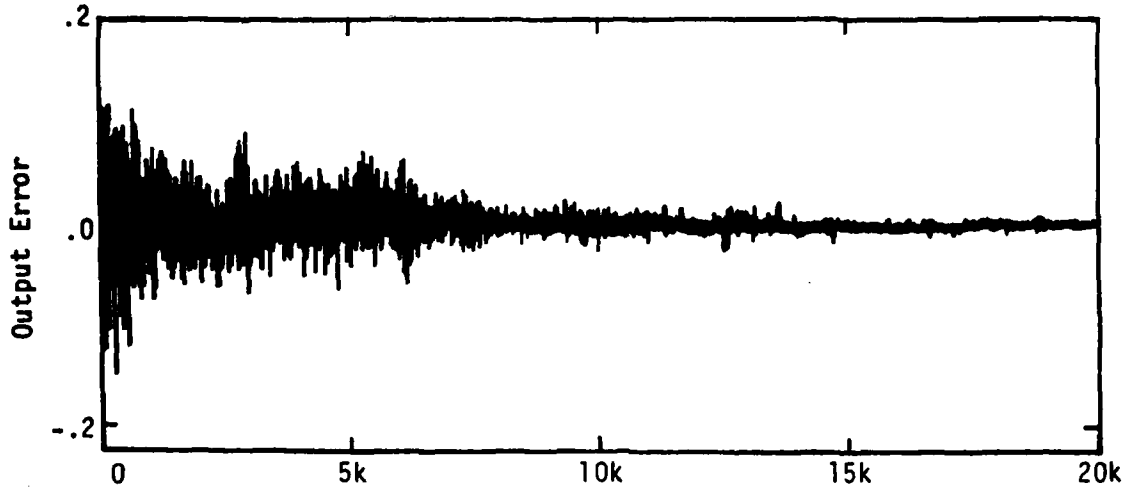
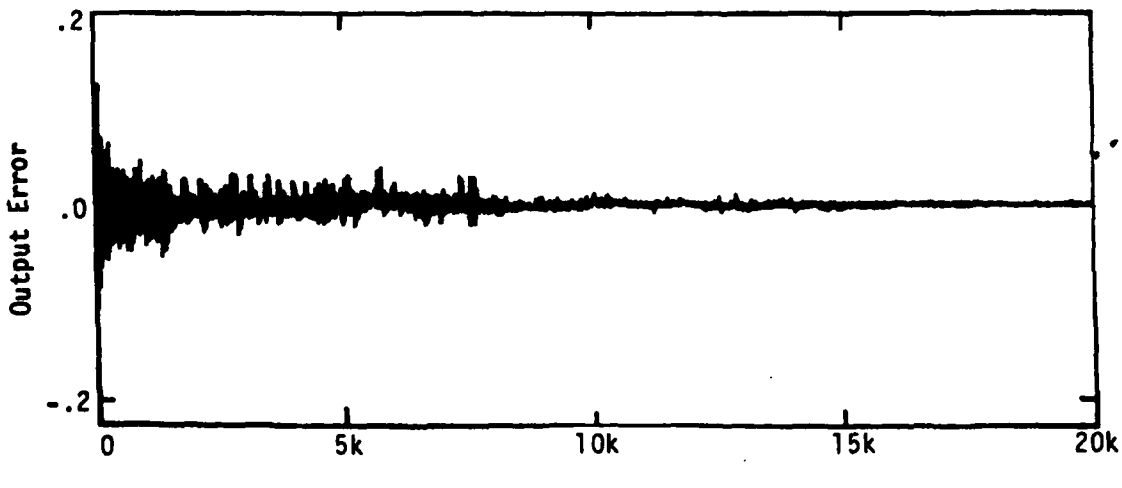


Figure 24. Modified Adaptive Cascade Structure.



a) $\ell = 4$



b) $\ell = 2$

Figure 25. Delay Adjustment Example.

o = actual zeroes

/ = actual poles

* = calculated zeroes

+ = calculated poles

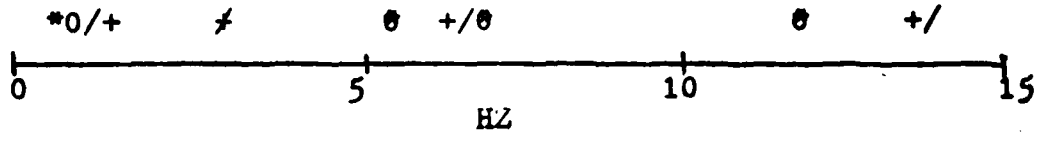


Figure 26. Adaptive ARMA pole-zero estimates compared to actual known values, after adapting to 3 data points.

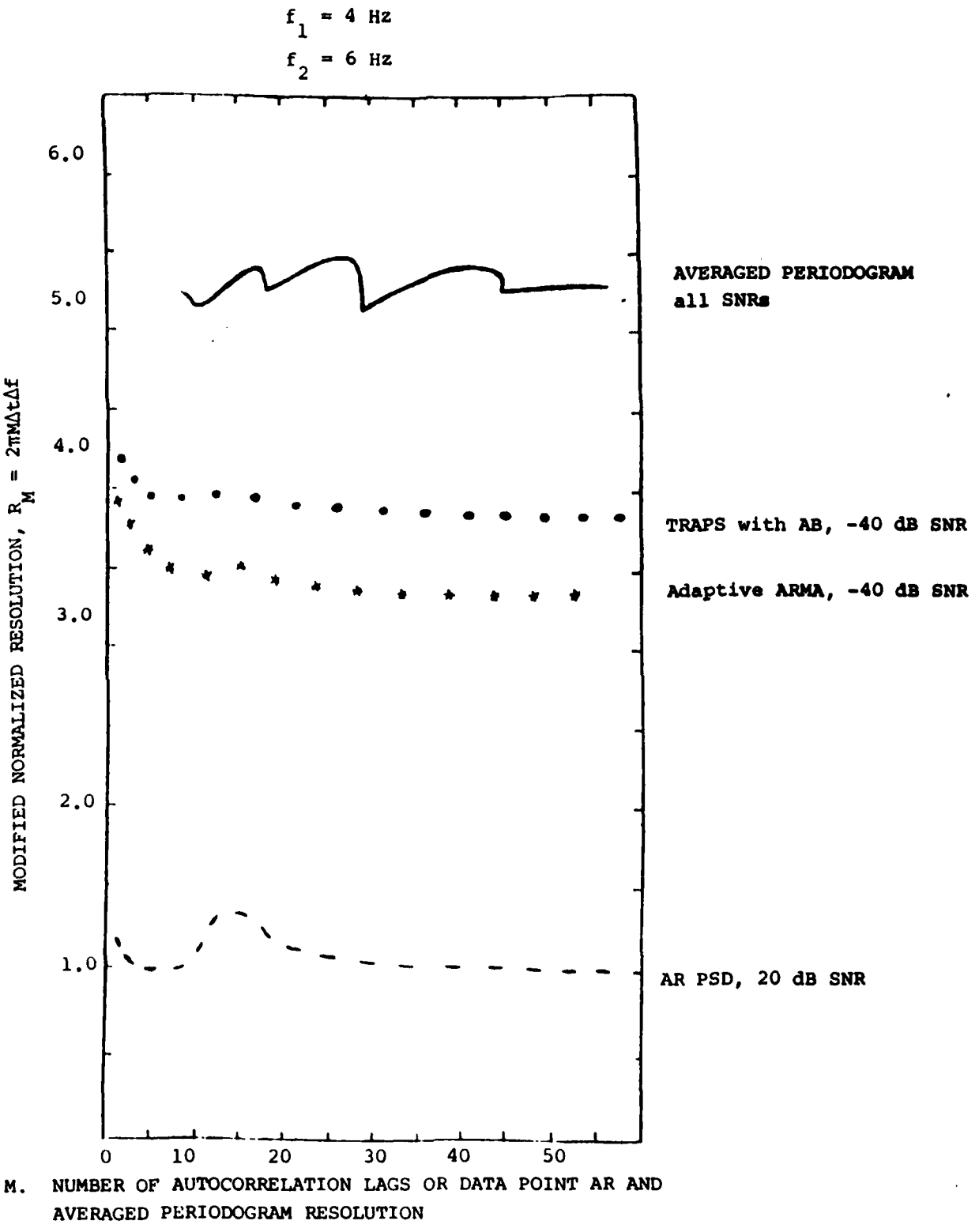


Figure 27. Results of Adaptive Autoregressive Moving Average (ARMA) Frequency Resolution Evaluation. Note that a smaller R denotes improved resolution.

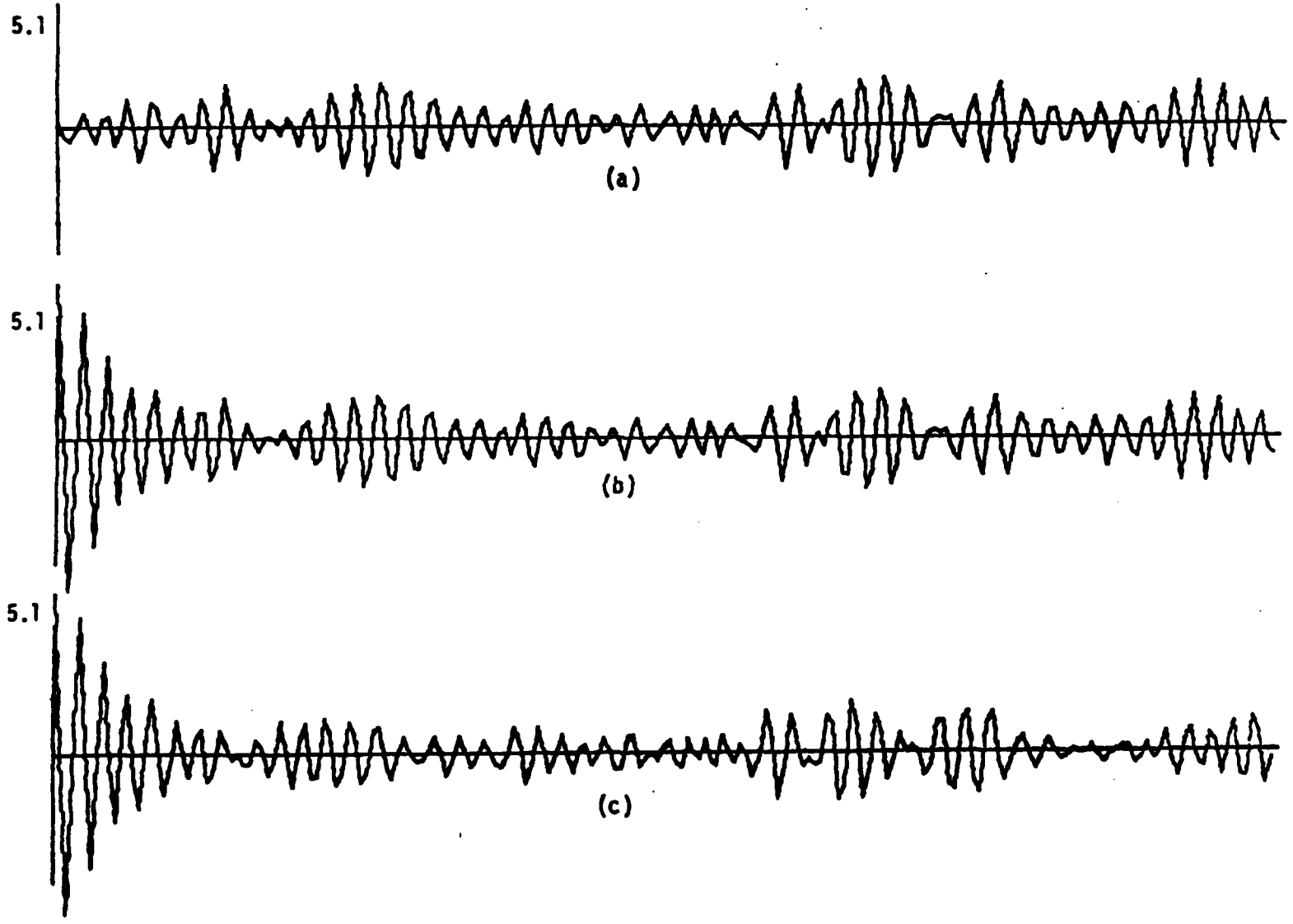


Figure 28. Comparison of time-domain outputs of TRAPS with and without slaved filter. (a) Signal input. (b) Slaved filter output. (c) TRAPS output.

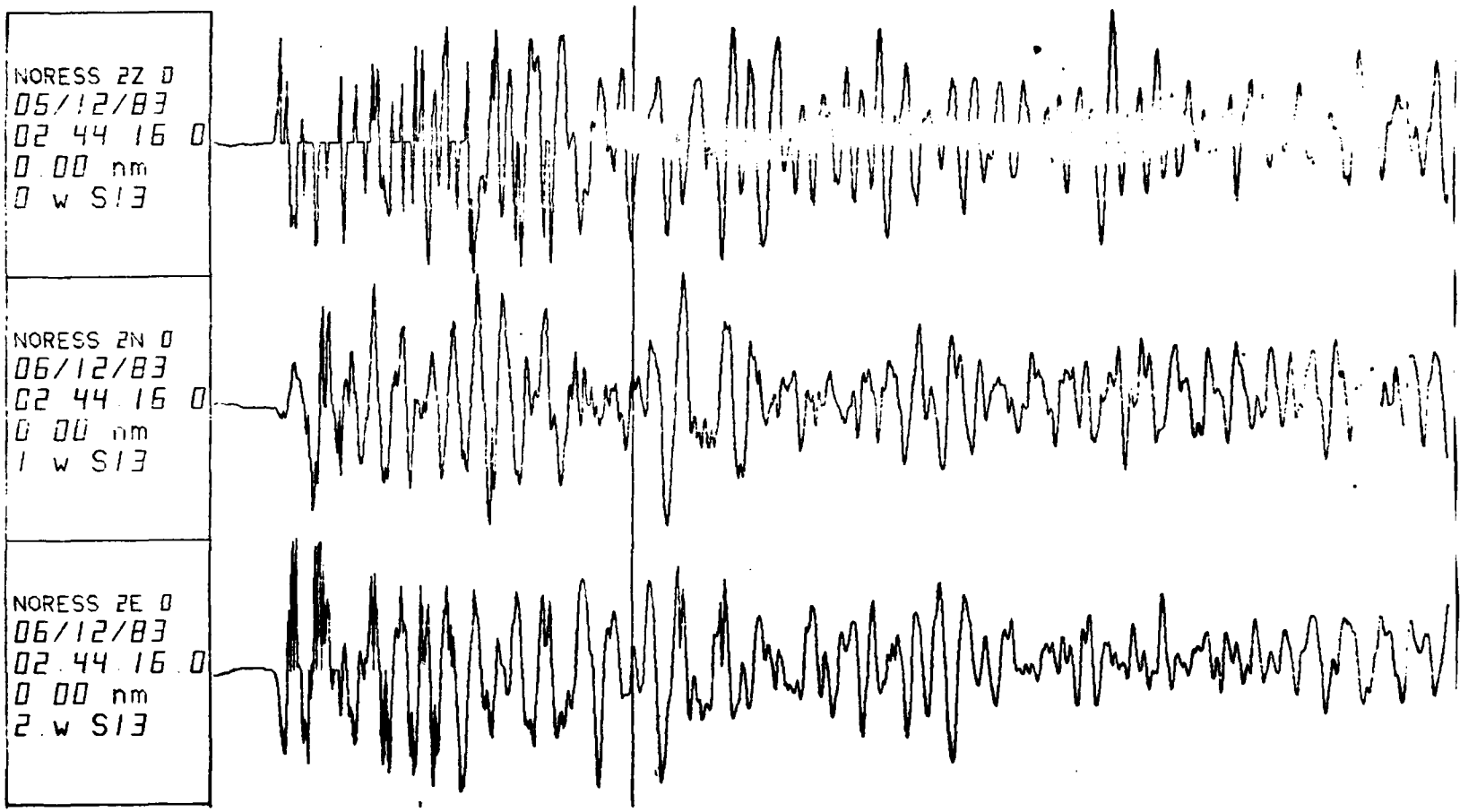


Figure 29a. Display of Event 1 Received Waveform

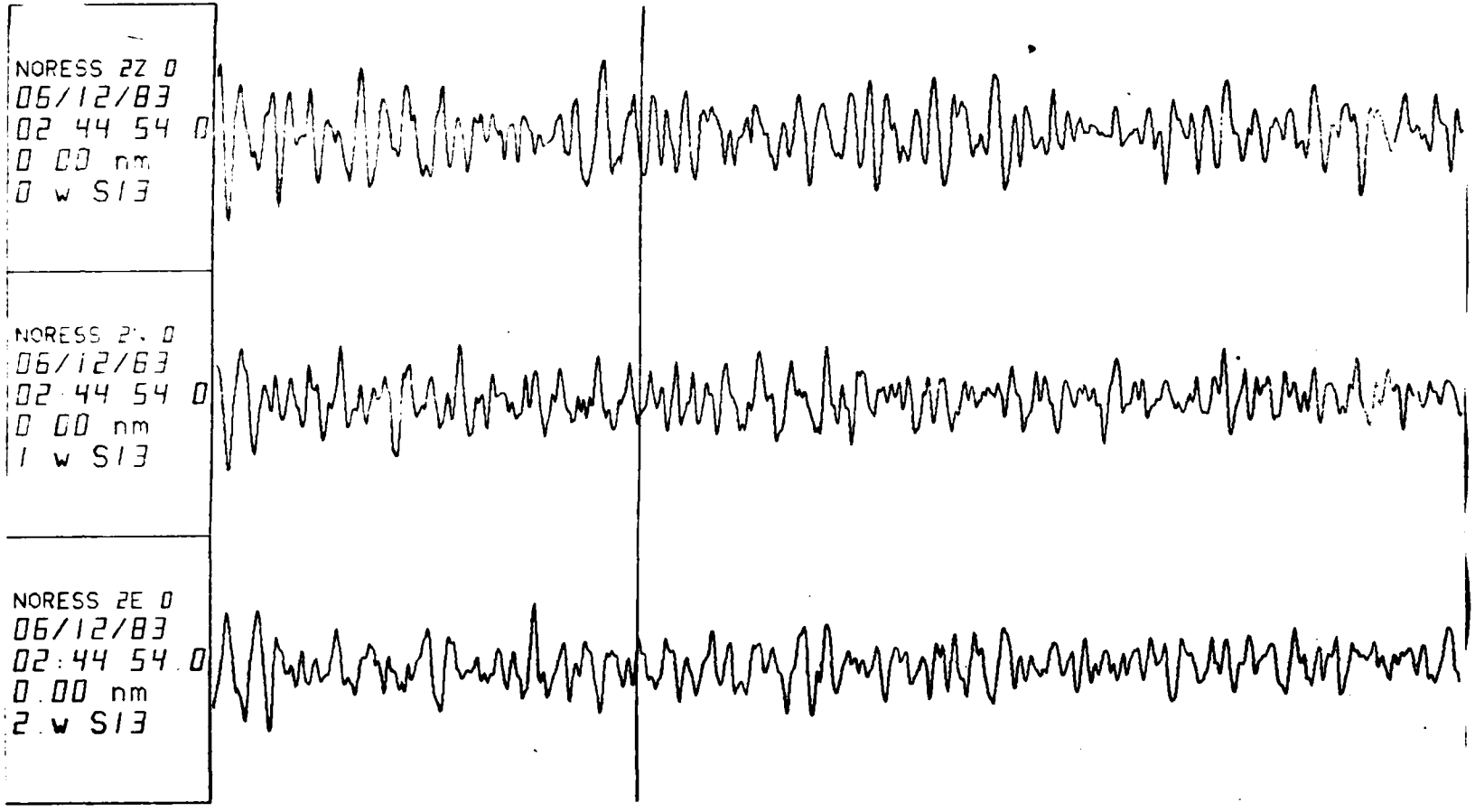


Figure 29b. Display of Event Received Waveform

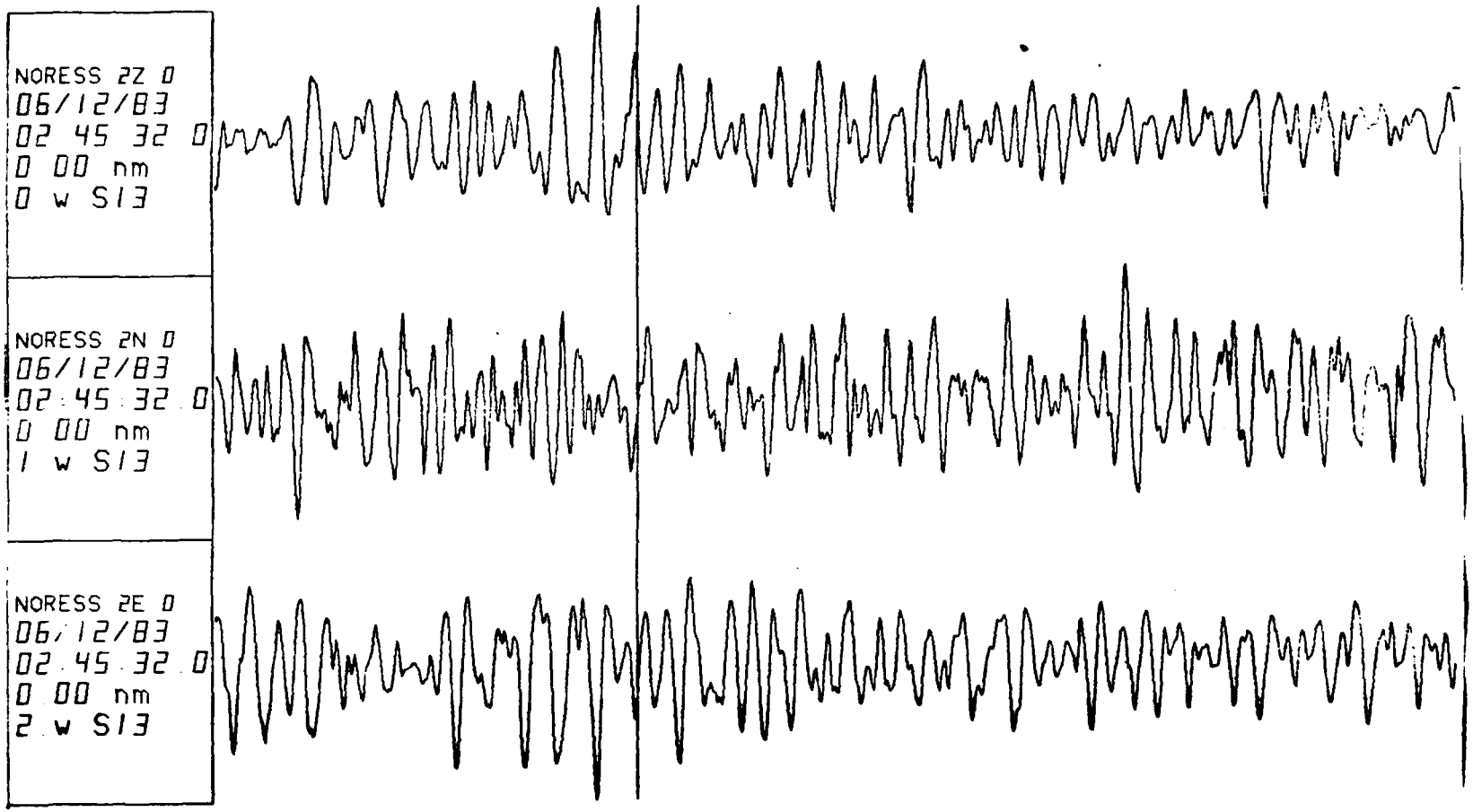


Figure 29c. Display of Event 1 Received Waveform

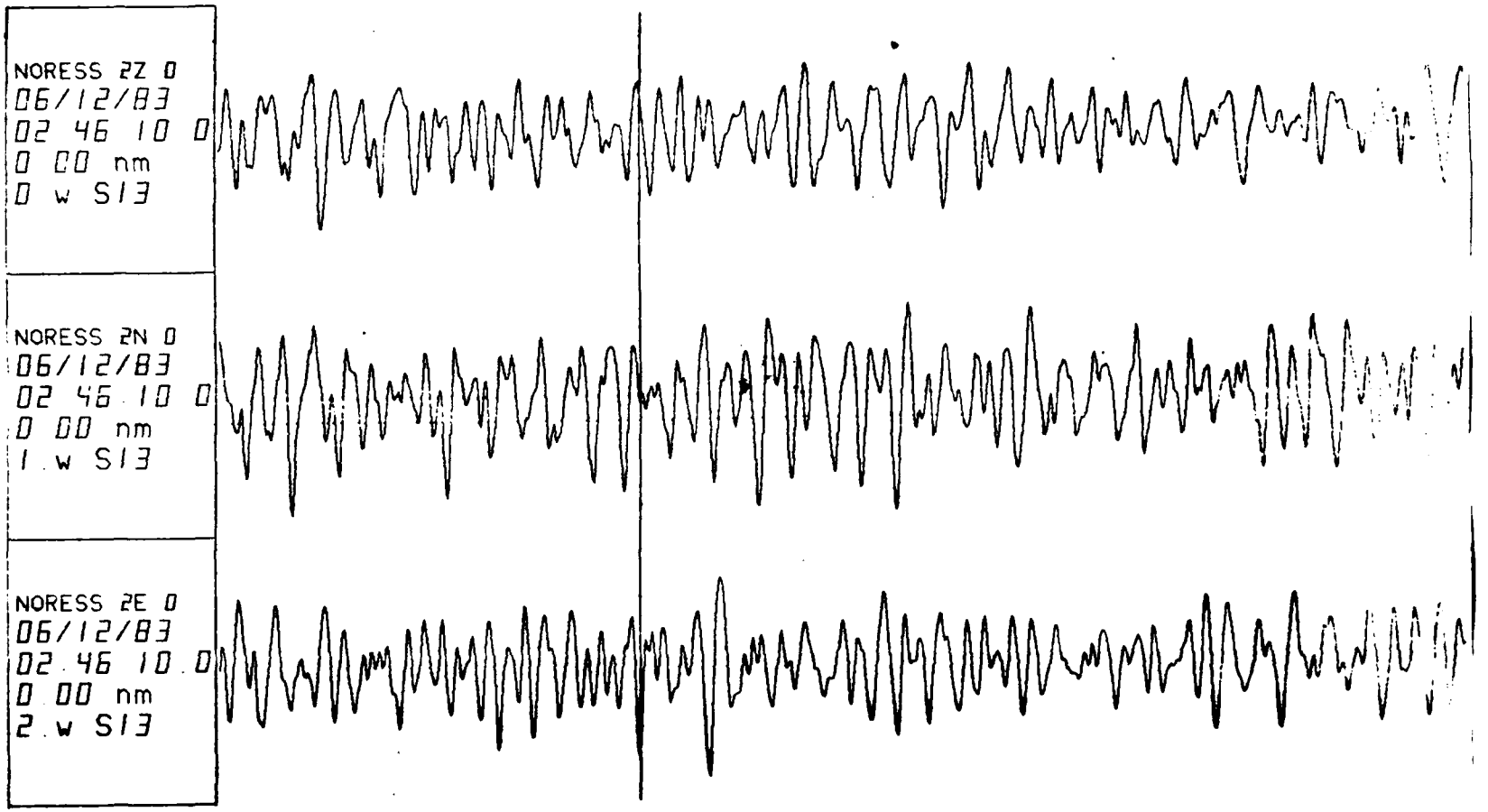


Figure 29d. Display of Event 1 Received Waveform

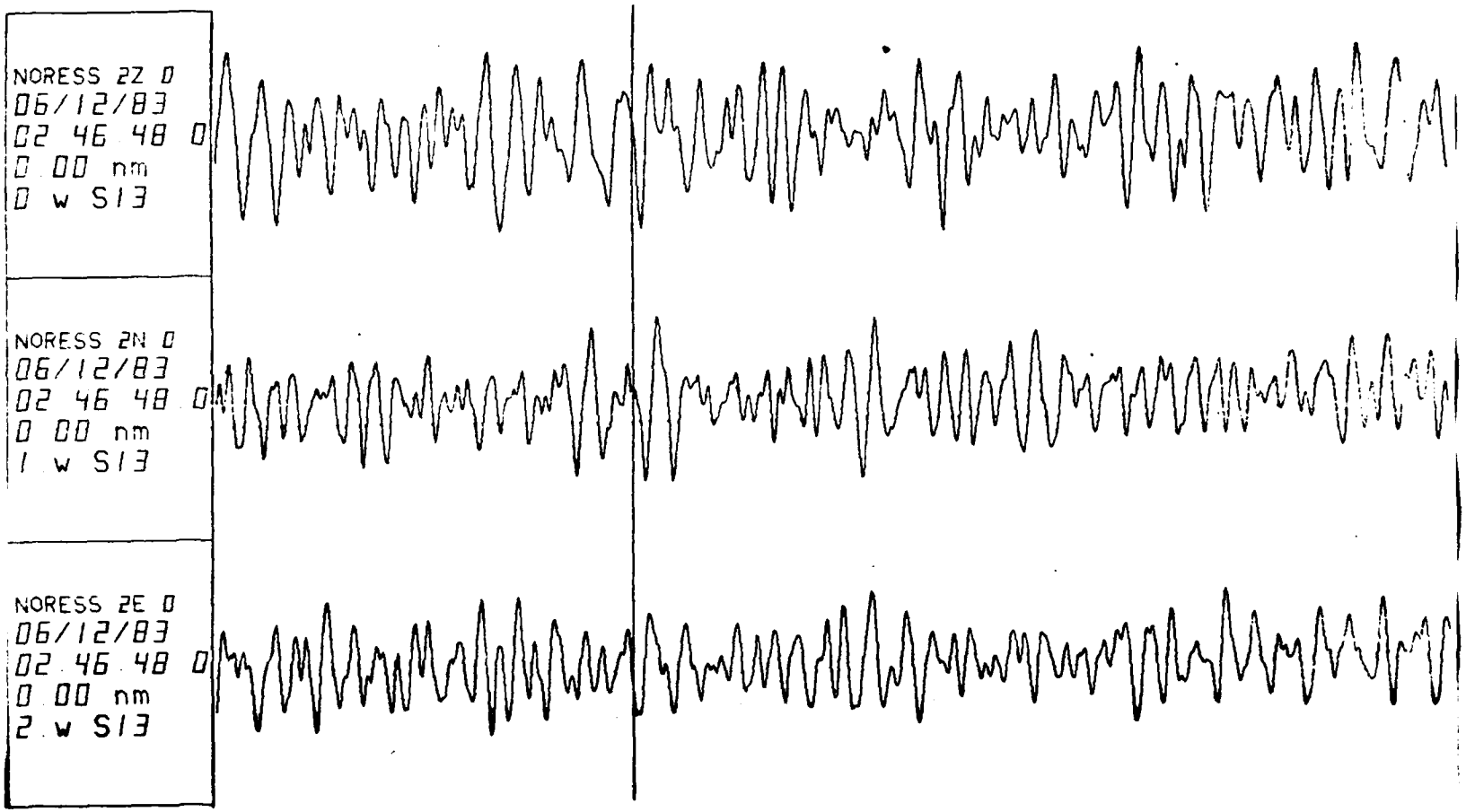


Figure 29e. Display of Event 1 Received Waveform

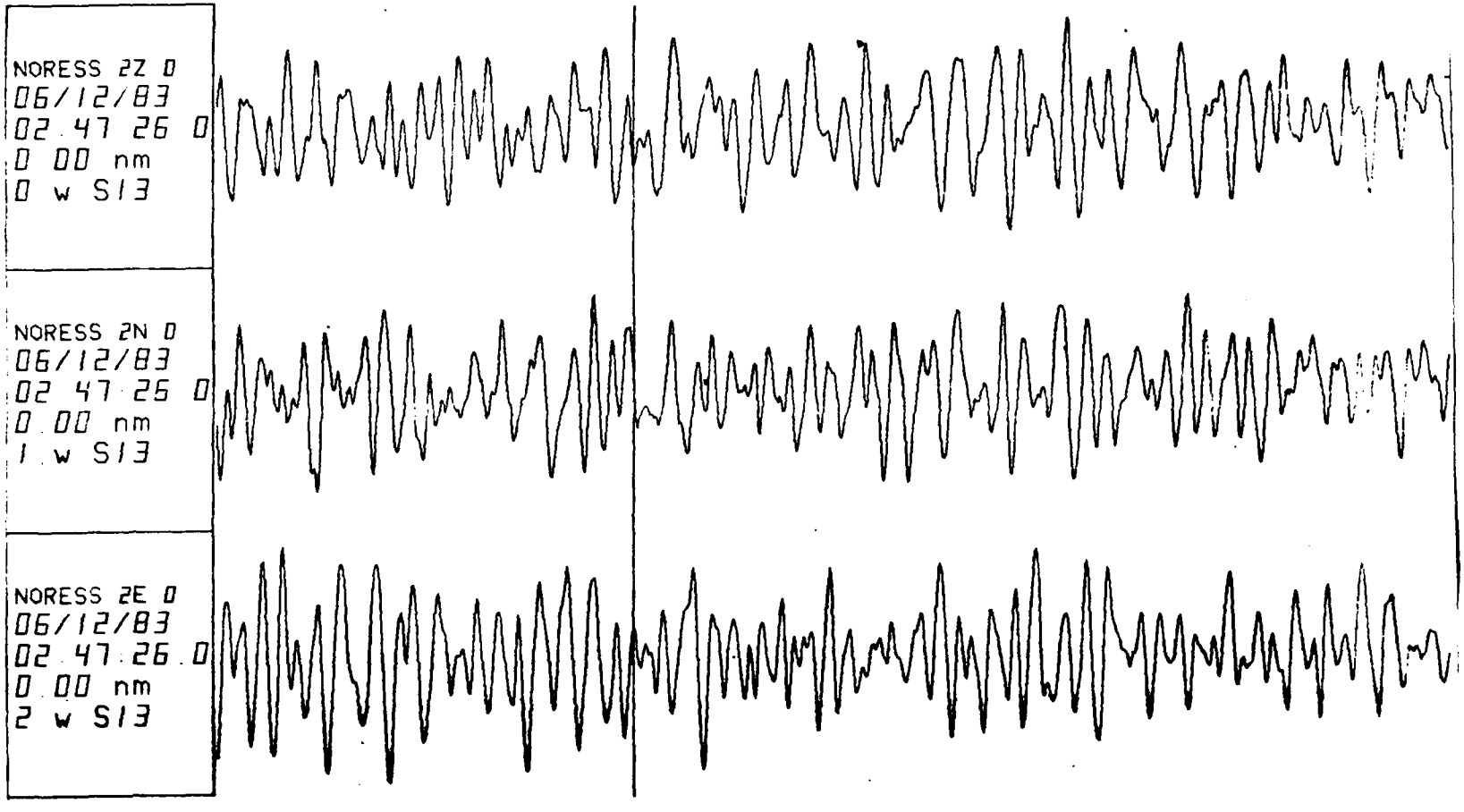


Figure 29f. Display of Event 1 Received Waveform

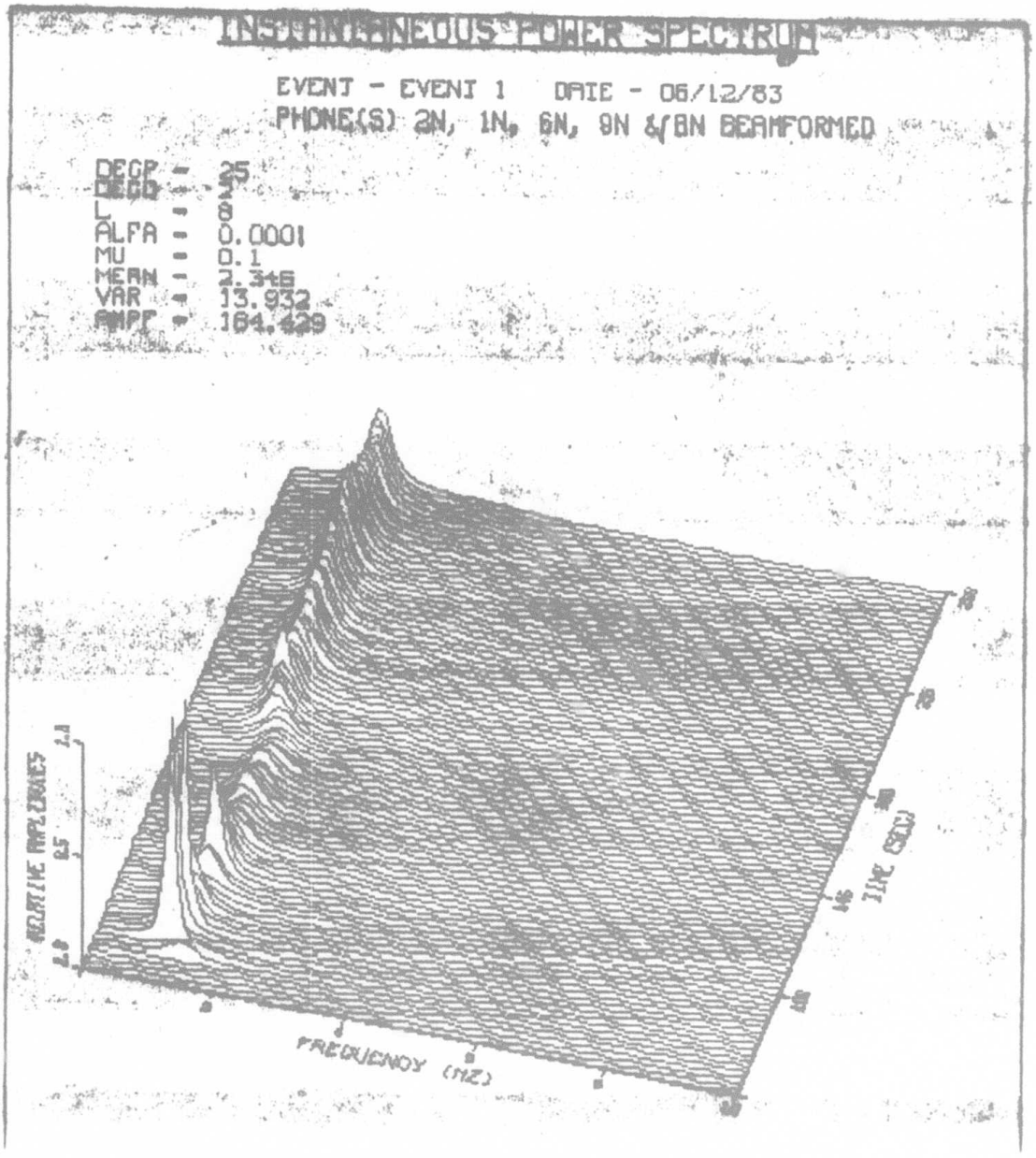

INSTANTANEOUS POWER SPECTRUM
EVENT - EVENT 1 DATE - 06/12/83
PHONE(S) 2N, 1N, 6N, 9N &/8N BEAMFORMED
DEGP = 25
DEGD = 2
L = 8
ALFA = 0.0001
MU = 0.1
MEAN = 2.346
VAR = 13.932
AMPT = 184.429
RELATIVE AMPLITUDES
FREQUENCY (HZ)
TIME (SEC)

Figure 30 a.

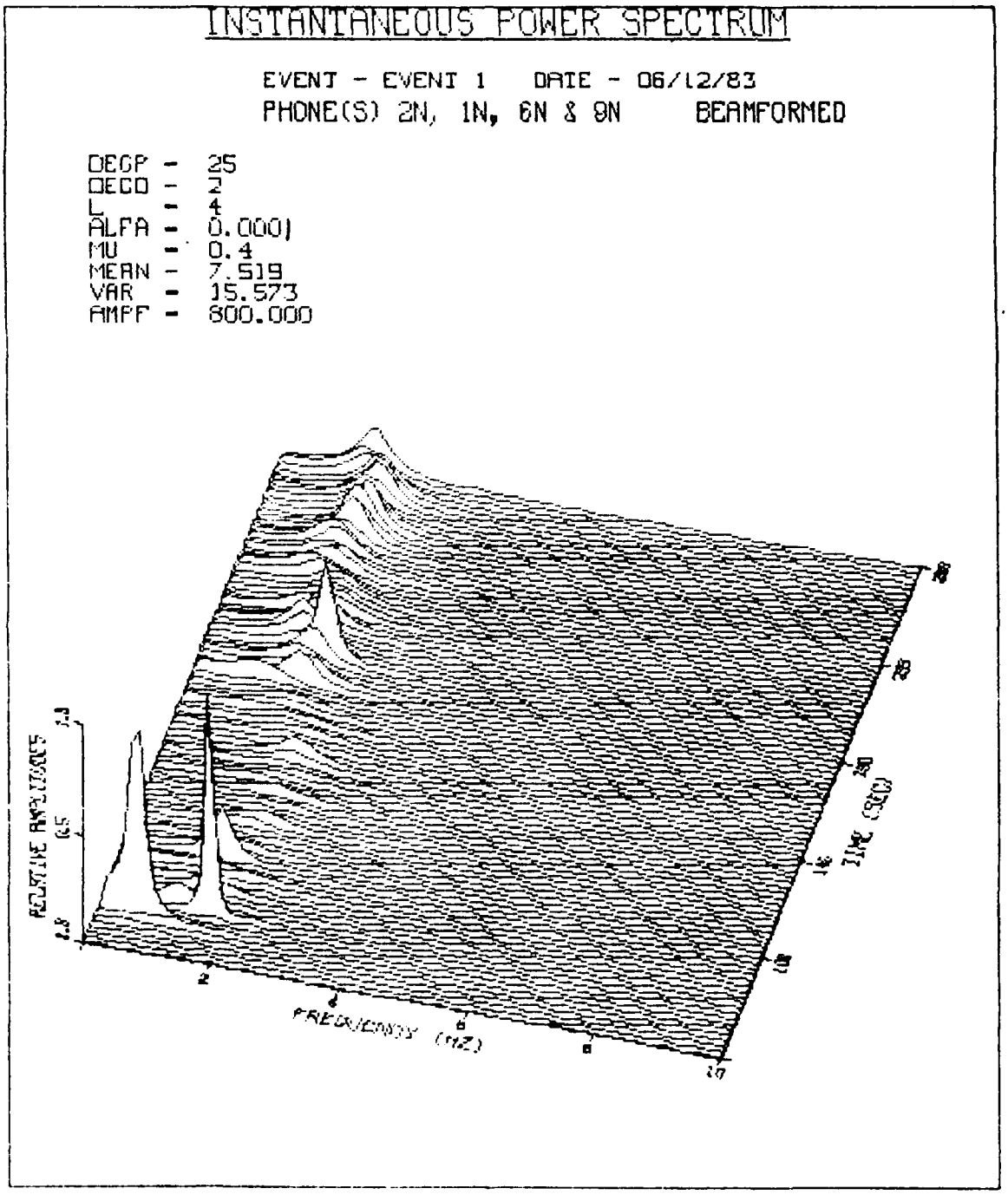

INSTANTANEOUS POWER SPECTRUM
EVENT - EVENT 1 DATE - 06/12/83
PHONE(S) 2N, 1N, 6N & 8N BEAMFORMED
DEGP - 25
DECD - 2
L - 4
ALFA - 0.0001
MU - 0.4
MEAN - 7.519
VAR - 15.573
AMPF - 800.000
RELATIVE AMPLITUDES
FREQUENCY (HZ)
TIME (SEC)

Figure 30b.

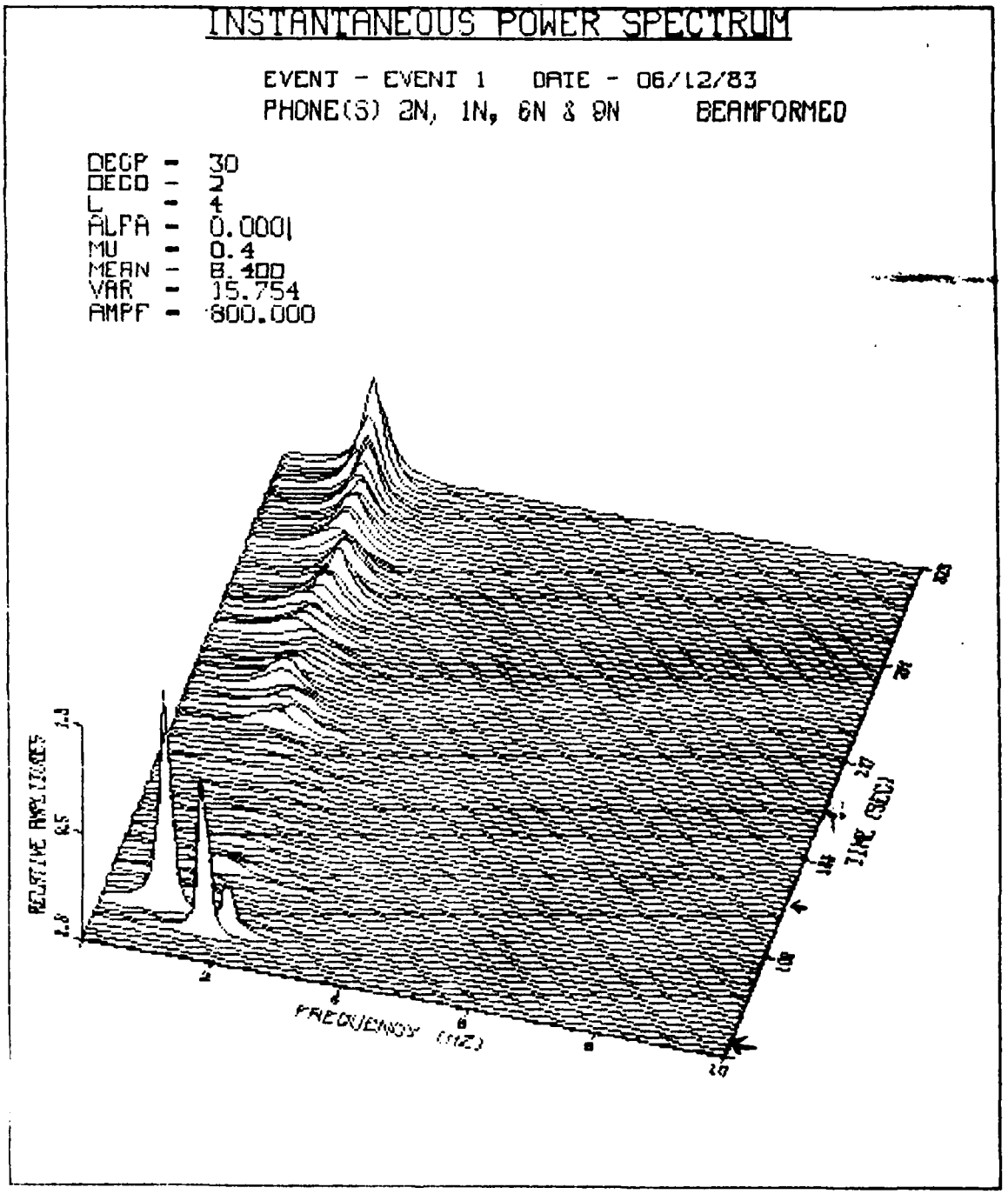

INSTANTANEOUS POWER SPECTRUM
EVENT - EVENT 1 DATE - 06/12/83
PHONE(S) 2N, 1N, 6N & 8N BEAMFORMED
DECP = 30
DECO = 2
L = 4
ALPA = 0.0001
MU = 0.4
MEAN = 8.400
VAR = 15.754
AMPF = 800.000
FREQUENCY (HZ)

Figure 30 c.

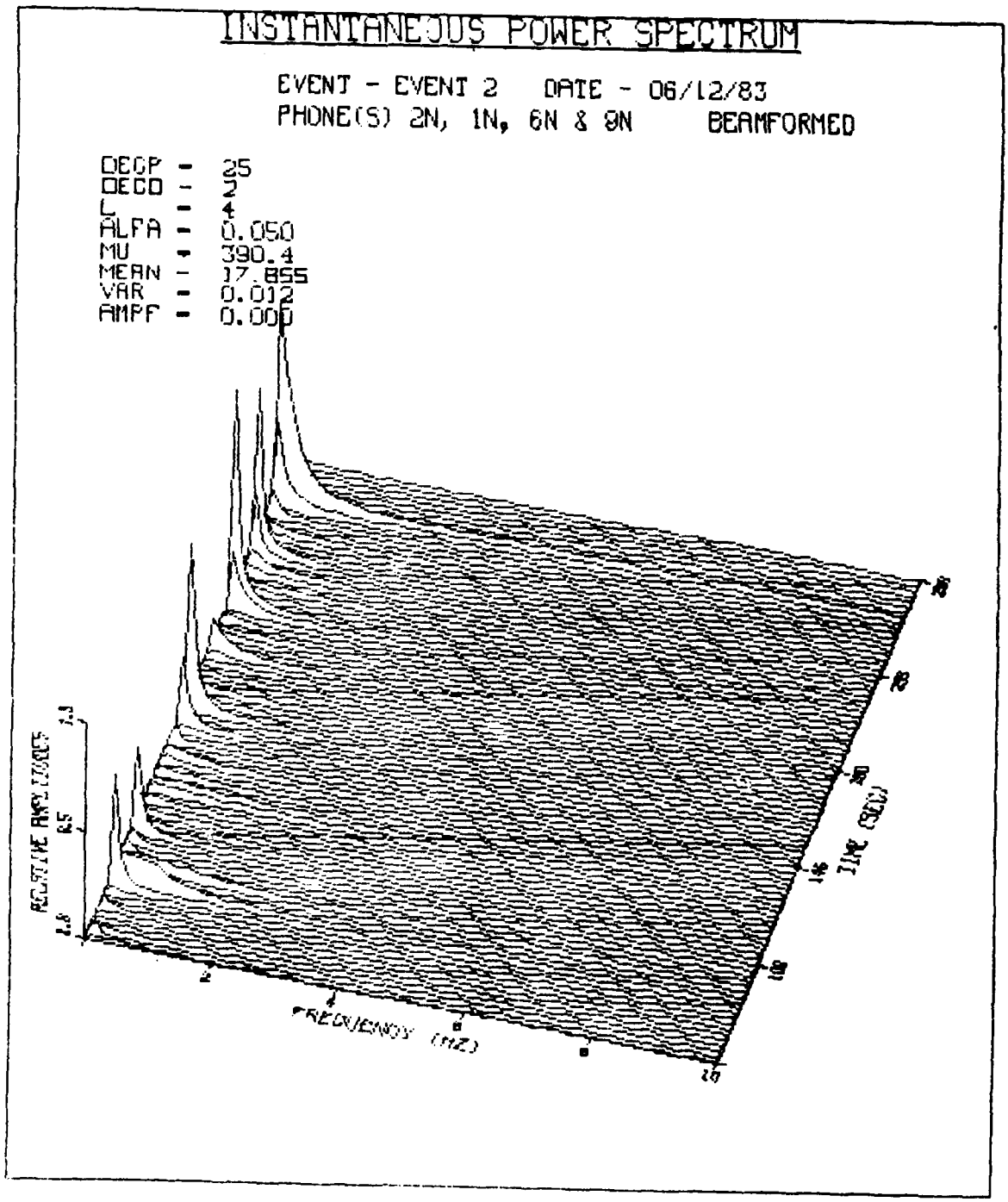

INSTANTANEOUS POWER SPECTRUM
EVENT - EVENT 2 DATE - 06/12/83
PHONE(S) 2N, 1N, 6N & 8N BEAMFORMED
DECP - 25
DECO - 2
L - 4
ALFA - 0.050
MU - 380.4
MEAN - 17.855
VAR - 0.012
AMPF - 0.000
FREQUENCY (HZ)
TIME (SEC)

Figure 30 d.

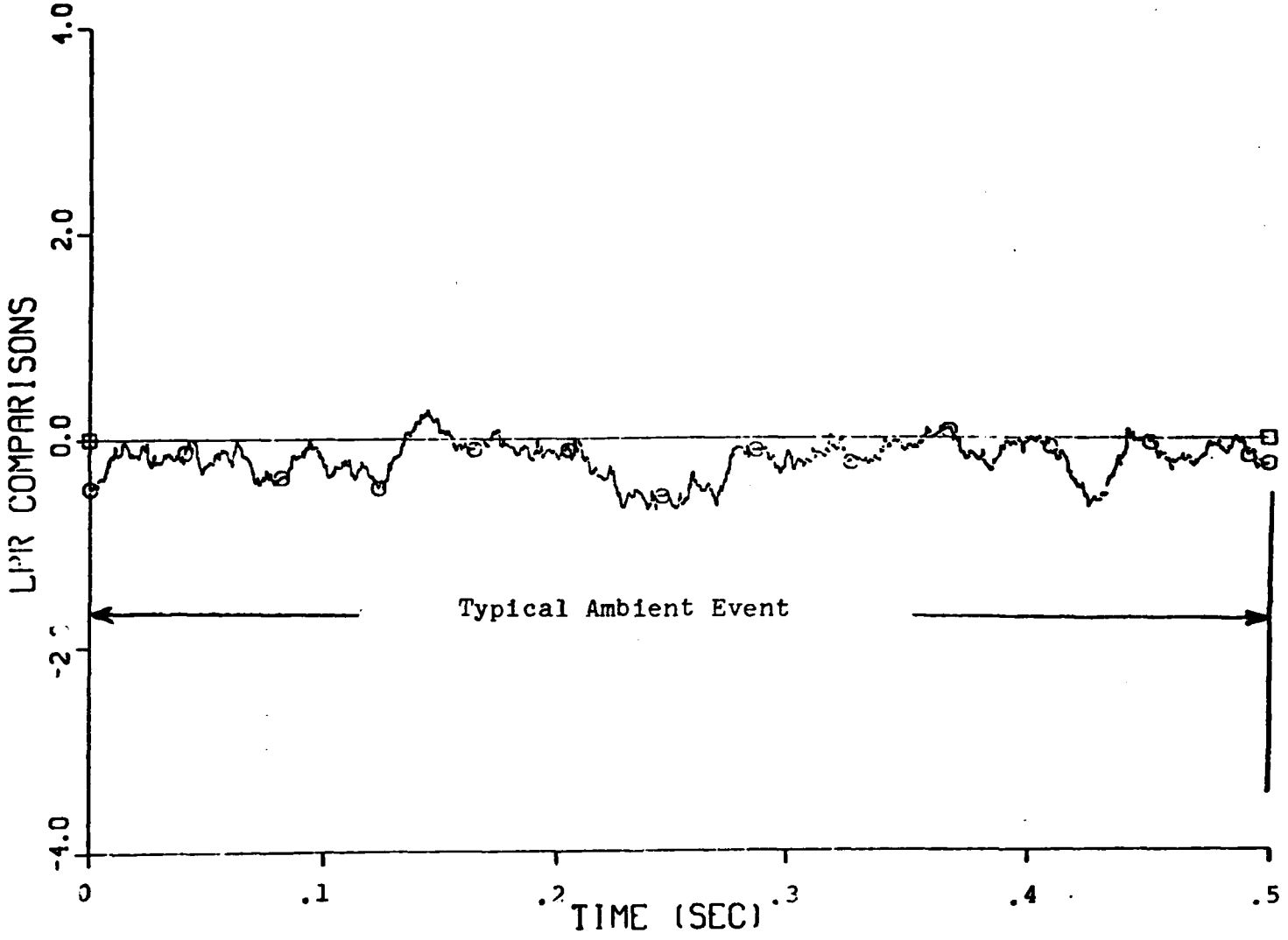


Figure 31. Example of TRAPS-LPR comparison of two ambient events

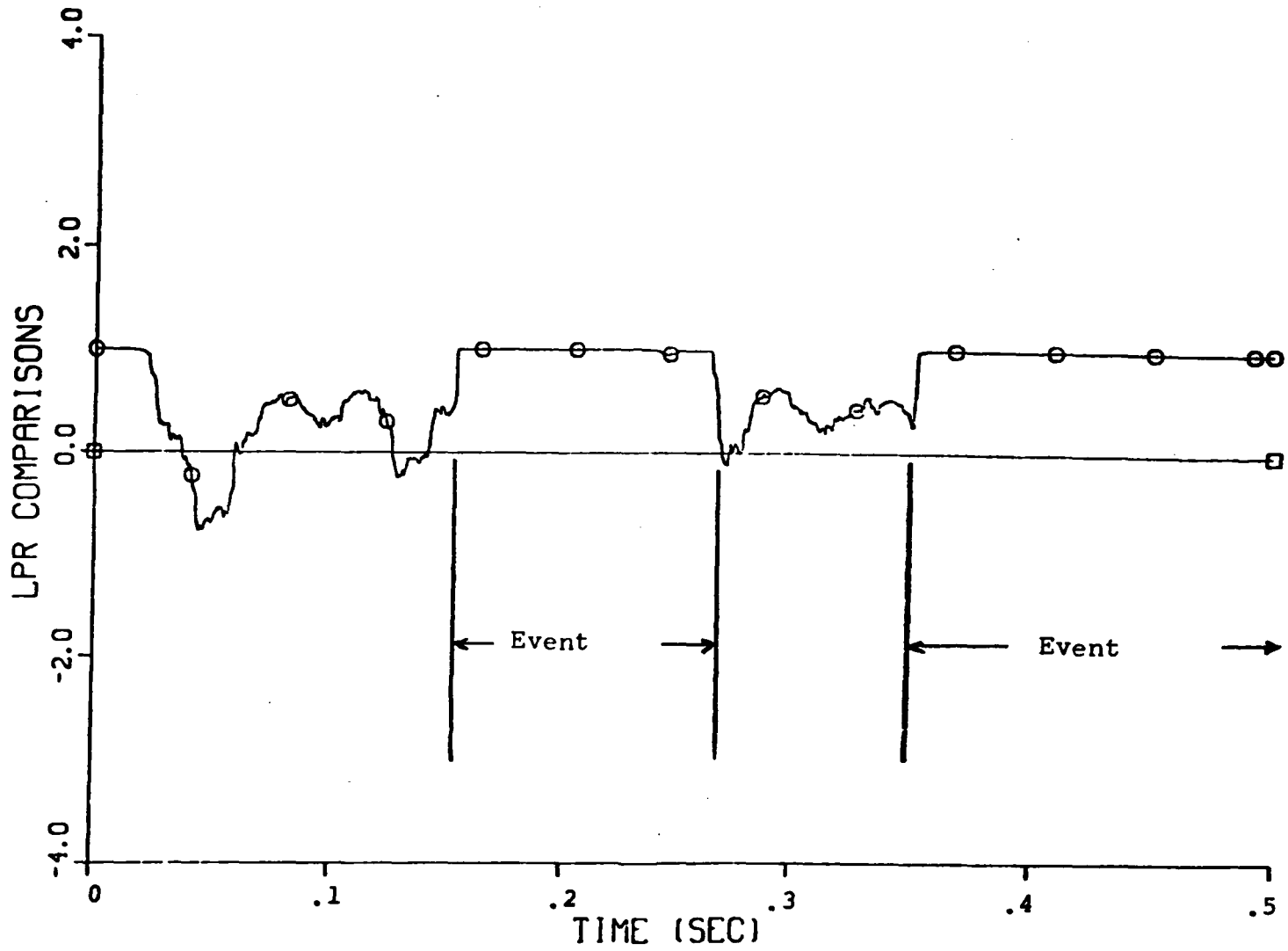


Figure 32. Example of TRAPS-LPR comparison of two similar types of events.

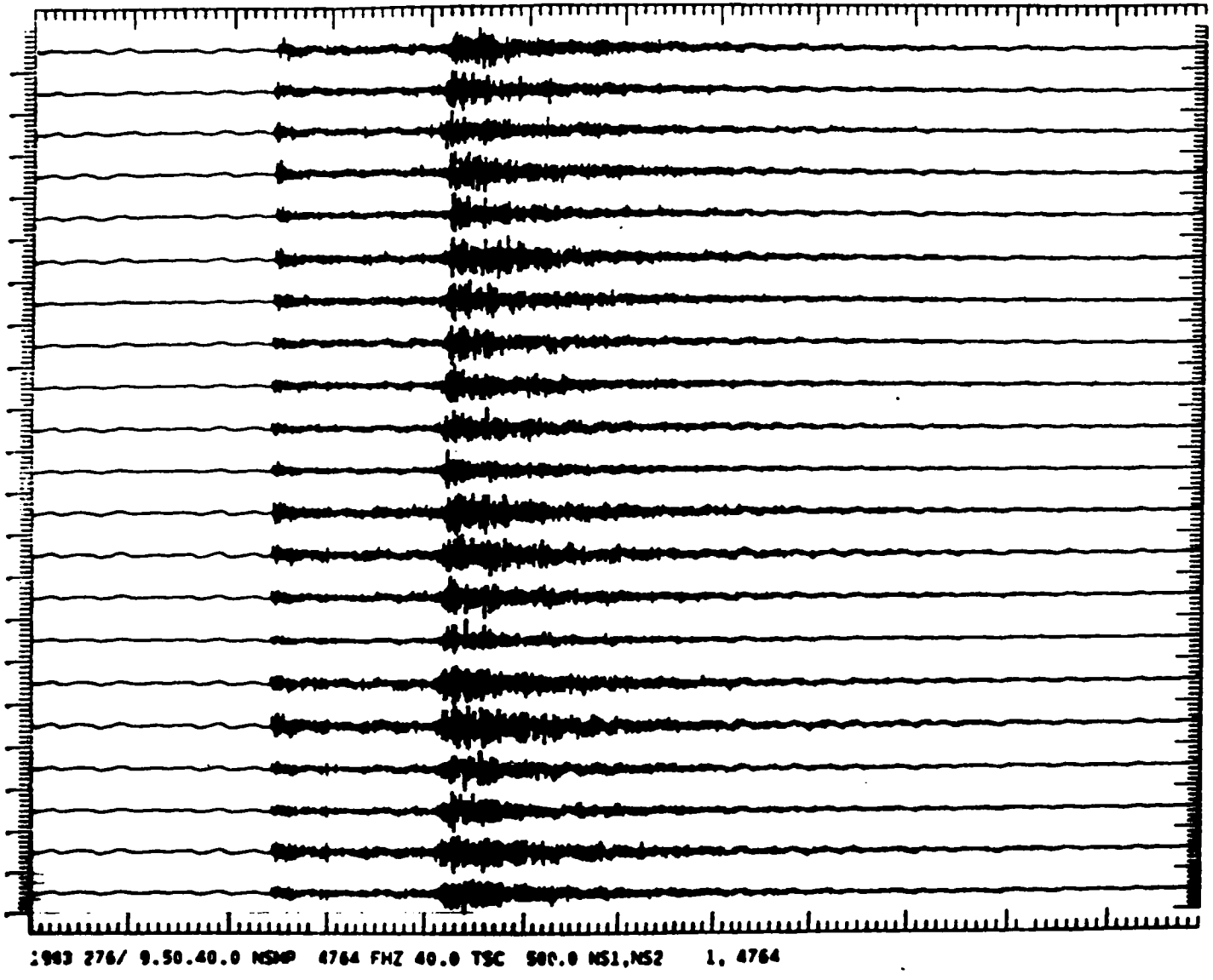


Figure 33. Event 2'

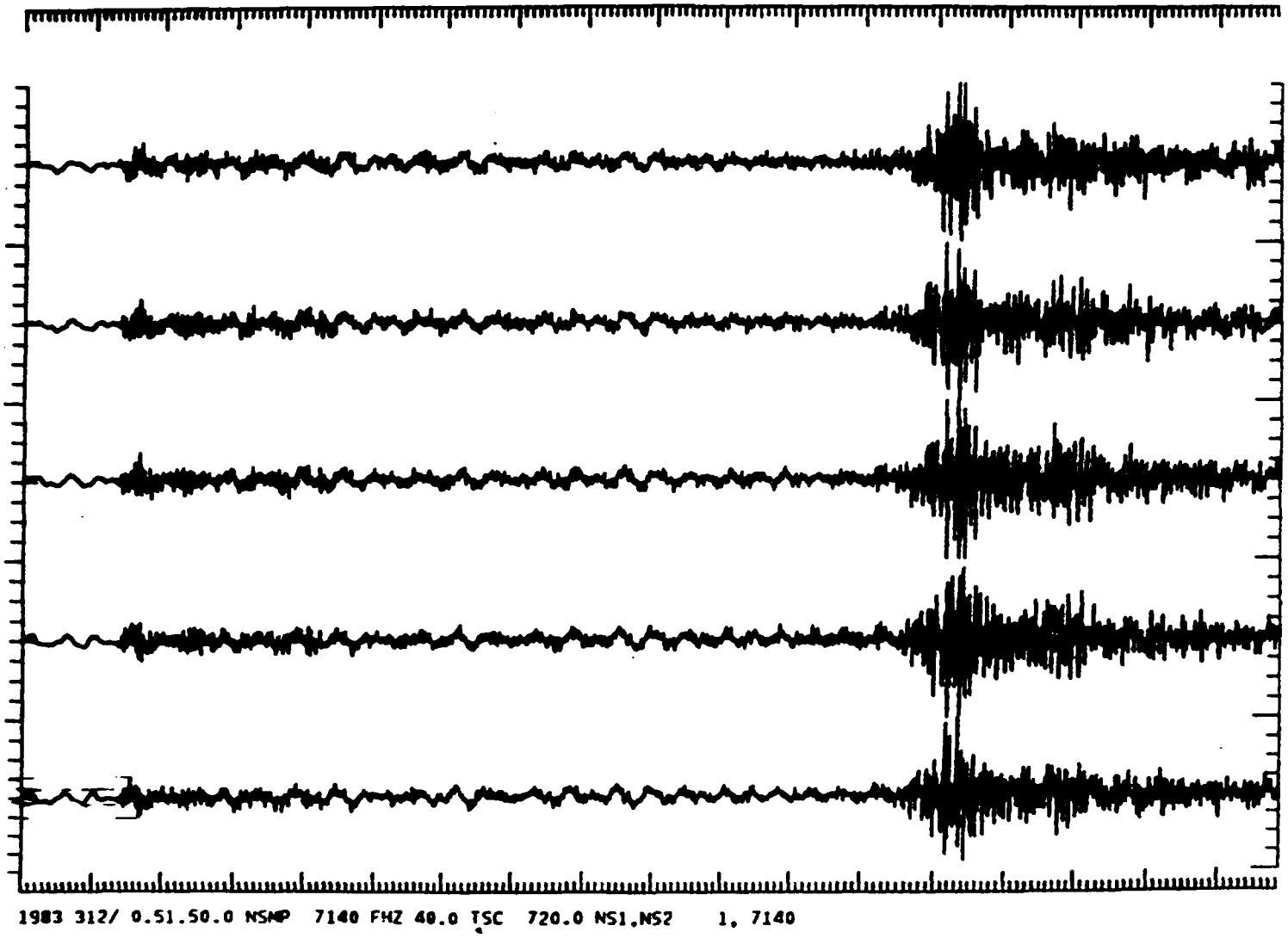


Figure 34. Event 8

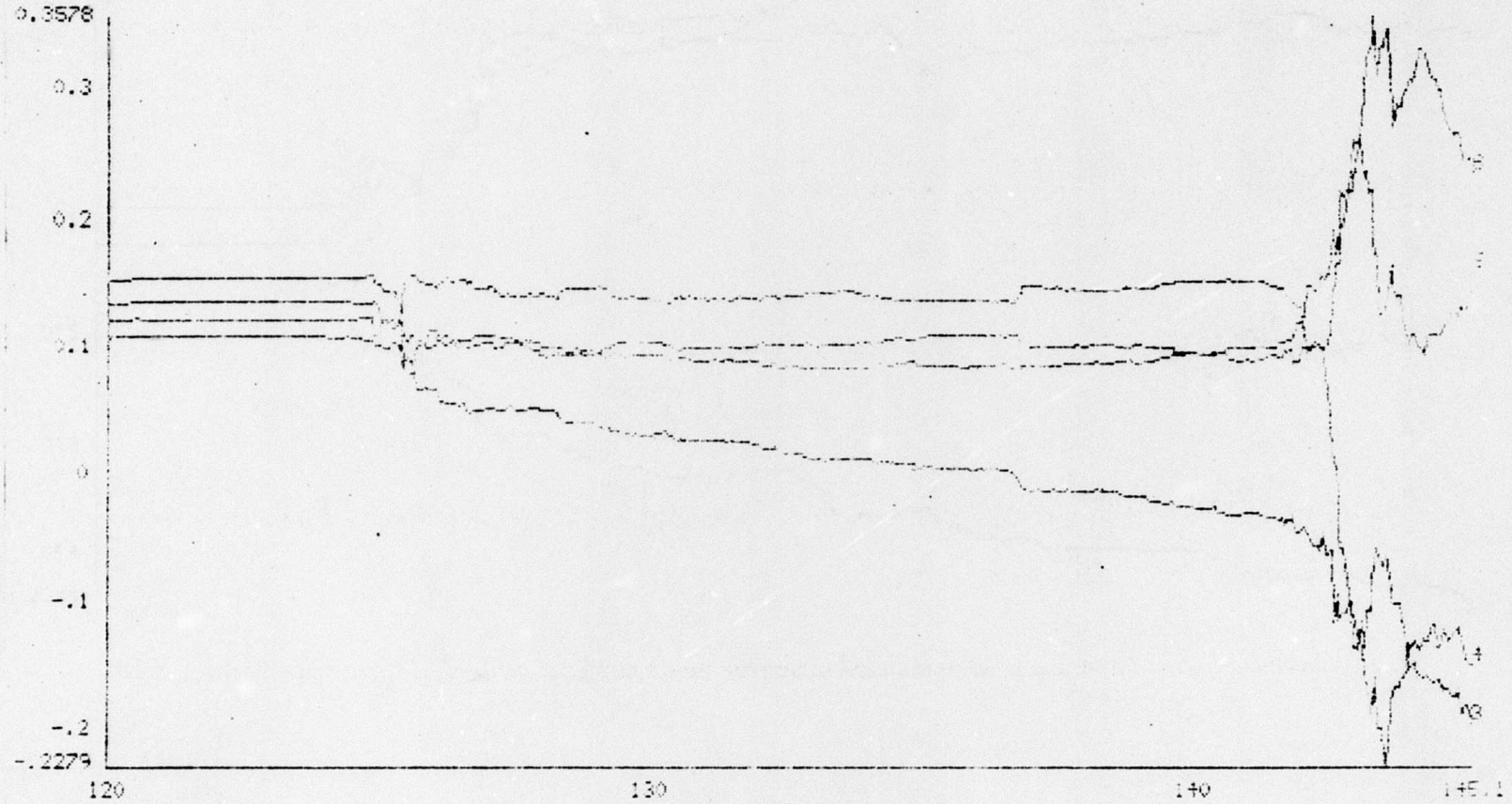


FIGURE 35. EVENT 2' COEFFICIENTS - 1ST WAVE

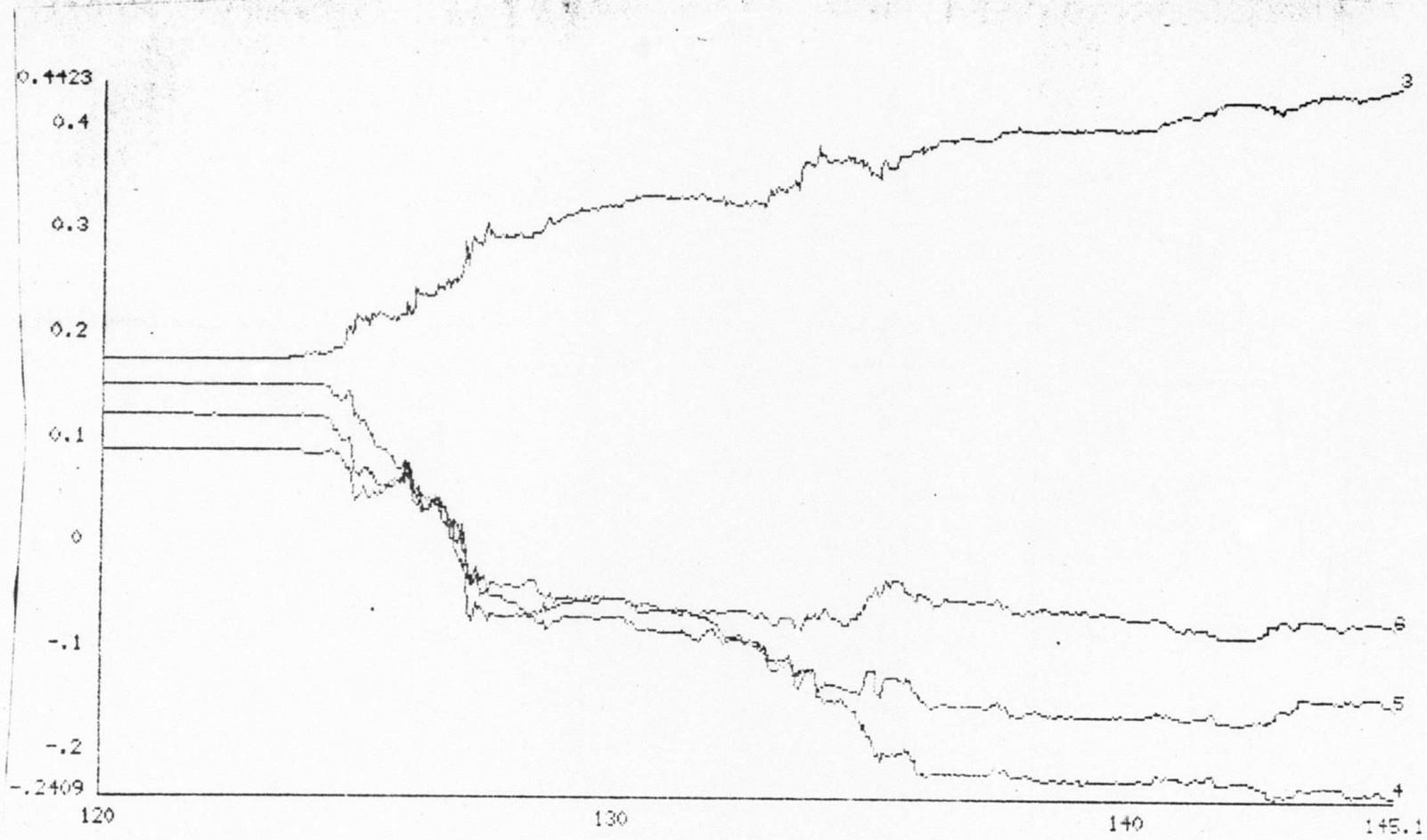


FIGURE 36. EVENT 8' COEFFICIENTS

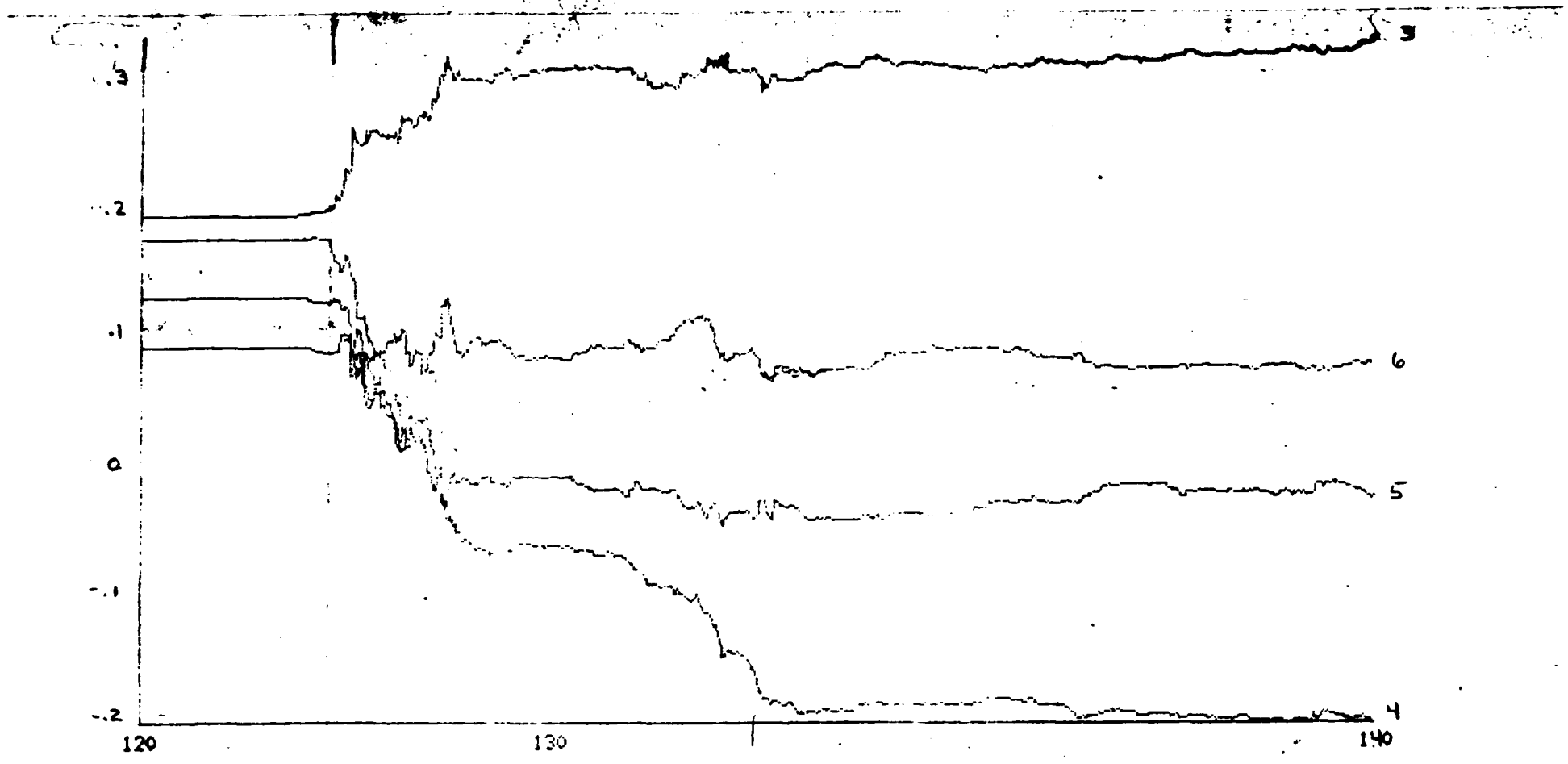


FIGURE 37. EVENT 2' WPC

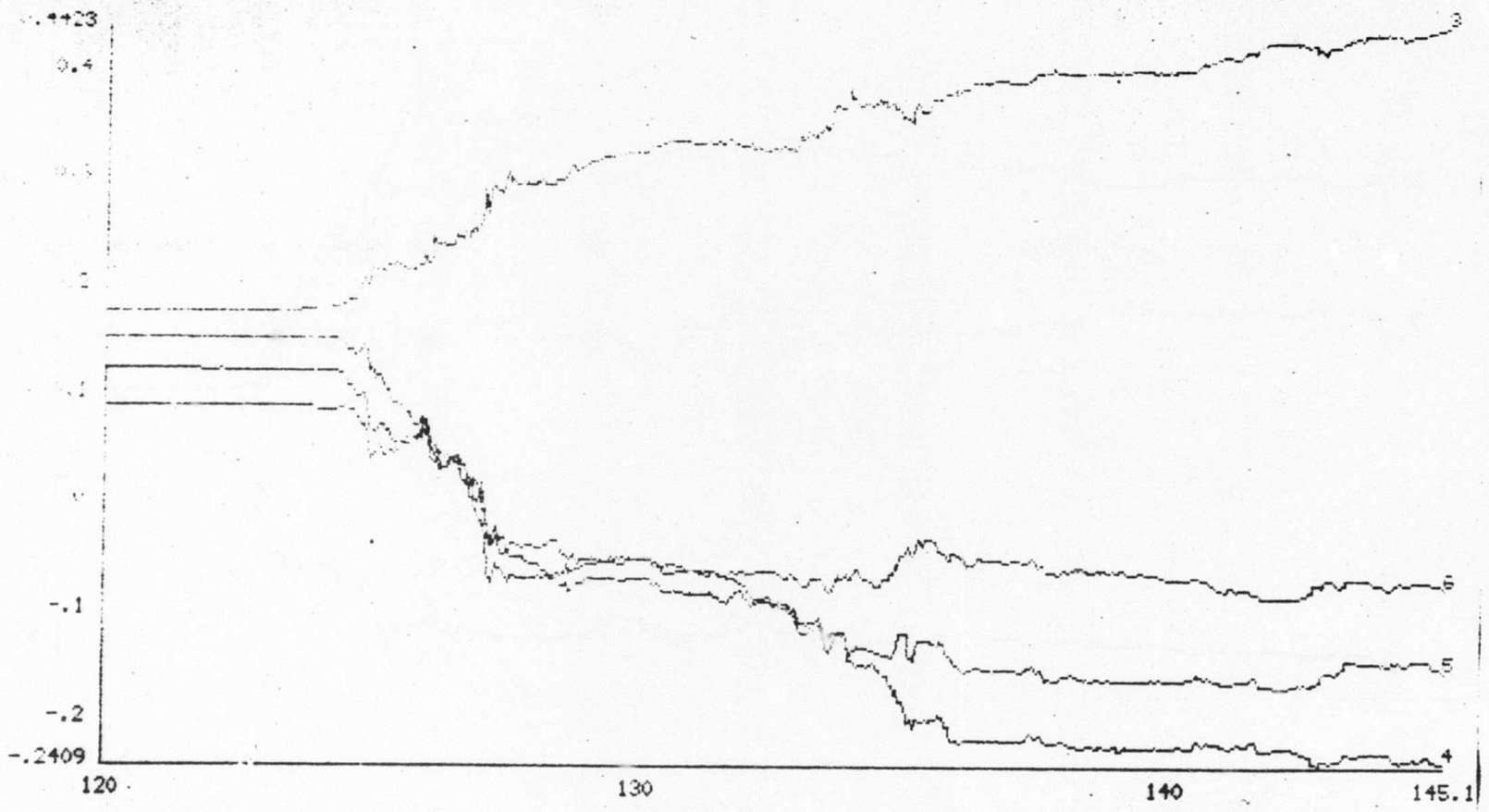


FIGURE 38. EVENT 8' WPC

APPENDIX A

APPENDIX A: 3-D Hidden Plot Program


```
      program plt3d
      dimension zmat (91, 91)
      dimension x(181), z(181), max (1500), param (9)
      data param /20.0, 30.0, 0.75, 0.0, 0.0, 0.0,
     1           6.5, 6.75, 2.0/
c
c... initialize variables
c
      deltax = param (1)
      deltay = param (2)
      deltaz = param (3)
      firstx = param (4)
      firsty = param (5)
      firstz = param (6)
      xlen   = param (7)
      ylen   = param (8)
      zlen   = param (9)
      yincr  = 1.0
      res = 40.
      n = 91
      rotate = 0.0
      angle = 65.0
      iter = 1
      do 10 i = 1, 1500
      max (i) = 0
 10   continue
      do 20 i = 1, n
      x (i) = float (i - 1)
 20   continue
      print '("reading data please wait")'
c
c...open input file
c
      open (7, file = 's102', status = 'old')
c
c...read all data
c
      read (7, 1001)
 1001 format (////).
      do 24 i = 1, 91
      ict = 0
      jct = 0
      do 22 j = 1, 13
      ict = jct + 1
      jct = ict + 6
      read (7, 1000) (zmat (i, K), K = ict, jct)
 22   continue
 24   continue
 1000 format (7(f7.5, 1x))
c
c...ring bell
c
      print '("")'
c
c...initiate plot
c
```

```
      call ginit
      call clipon
      call uvwind (237., 4095., 0., 3119.)
      call xywind (0.0, 8.5, 0.0, 6.75, 0, 0)
c
c...clear screen
c
      call erase
c
c...calculate shifts of origin
c
      rangle = angle * 3.1415927 / 180.0
      yinc = yincr / deltay
      zshift = yinc
      xshift = yinc * cos (rangle)
c
c...rotate plot
c
      alph = rotate * 3.141527 / 180.0
      sina = sin (alph)
      cosa = cos (alph)
c
c...main loop to plot
c
      do 120 k = 1, n
c
c...take data one line at a time
c
      iflag = 1
      do 110 j = 1, n
      z(j) = zmat (j, k)
      x1 = (x(j) - firstx) / deltax + float (iter - 1) * xshift
      z1 = (z(j) - firstz) / deltaz + float (iter - 1) * zshift
      if (j .ne. 1) go to 30
      x2 = x1
      z2 = z1
      x2p = cosa * x2 - sina * z2
      z2p = cosa * z2 + sina * x2
      x9 = x2p
      z9 = z2p
      call xymove (x9, z9)
c
c...interpolate, pick integral endpoints
c
      slope = 0.0
      go to 40
 30   continue
      slope = (z1-z2) / (x1-x2)
 40   continue
      b = z1 - (slope * x1)
      nstart = int (x2 * res)
      nend = int (x1 * res)
      x2p = cosa * x2 - sina * z2
      z2p = cosa * z2 + sina * x2
c
c...plot at points corresponding to values in max array,
```

```
c...as well as x2, if visible
c
      do 100 i = nstart, nend
      x3 = float (i) / res
      z3 = slope * x3 + b
      x3p = cosa * x3 - sina * z3
      z3p = cosa * z3 + sina * x3
      iz4 = int (z3 * 400.0)
      if (nstart .ne. i .or. i .eq. 1) go to 60
      iz2hat = max (i) - int (res * (max (i) - max (i-1)) * (x3-x2))
      iz5 = int (z2 * 400.0)
      if (iz5 .lt. iz2hat) go to 50
      call xydraw (x2p, z2p)
      iflag = 0
      go to 60
 50   continue
      call xymove (x2p, z2p)
      iflag = 1
 60   continue
      if (iz4 .lt. max (i)) go to 80
      if (iflag .eq. 1) go to 70
      call xydraw (x3p, z3p)
      max (i) = iz4
      go to 100
 70   continue
      iflag = 0
      go to 90
 80   continue
      iflag = 1
 90   continue
      call xymove (x3p, z3p)
 100  continue
c
c...update x2, z2, iteration counter
c
      z2 = z1
      x2 = x1
 110  continue
      iter = iter + 1
 120  continue
c
c...alpha mode and ring bell
c
      call xymove (0.0, 6.75)
      call size (2)
      call alpha
      print '(^G)'
c
      stop
      end
```

APPENDIX B

APPENDIX A

Stage Three of TRAPS Filter: Range and Bearing Estimation

The adaptive filter accurately determined the existence of the signal within some time interval. This time interval was then used as an input for time delay estimation. The exactness of the adaptive filter in specifying the signal permitted the time delay procedure to reliably specify the relative arrival times at each acoustic sensor. Once the arrival times of the signal at each sensor were estimated, a maximum likelihood method was able to estimate the range and bearing of the source. We will detail the mathematics at this time.

As briefed in the above paragraph, the first two stages of TRAPS allows one to stipulate the condition of a signal at all geophones within some interval of time (T_1, T_2). We can then estimate their relative arrival times by pairwise cross correlation of a filtered series measured at each geophone with that of a common, arbitrarily chosen, reference phone. Let $Y_o(t)$ be the filtered time series measured at the reference geophone, and let $Y_1(t)$ denote the filtered series on the second phone. Then

$$Y_o(t) = S(t) + n_o(t)$$

and

$$Y_1(t) = S(t-\tau) + n_1(t)$$

where τ (possibly negative) is the time by which the received signal at geophone 1 lags that at geophone 0, and $S(t)$ denotes the unknown filtered signal waveform. We assume only that this "signal" lies within the interval. More precisely, the filtered wave form is assumed to be similar on each geophone; some dissimilar residue in the signal may be combined with the noise so

long as it does not greatly affect our assumptions regarding the noise. We assume the noise $n_o(t)$ and $n_1(t)$ to be stationary, to be uncorrelated, and to be zero mean Gaussian processes with variances $\sigma^2_{\eta_o}$ and $\sigma^2_{\eta_1}$ respectively.

We choose the delay τ which maximizes the a postori probability density of τ given the measured series y_o, y_1. That is, we choose τ such that $p(\tau/y_o, y_1)$ is maximized.

It can be shown (14) by this criterion that $\Delta = t - t^*$, the estimated minus the true delay, has the standard deviation

$$\sigma_\Delta = \frac{\sqrt{8}}{2\ W} \left[\frac{2E}{N}\right]^{-\frac{1}{2}}$$

where W is the rms bandwidth in Hz and $\frac{2E}{N}$ is twice the signal energy to noise energy ratio. For the NORESS data listed in Table 1, we achieved an average standard deviation of

$$\sigma_{\Delta_{AVE}} = .018 \text{ sec.}$$

This cross correlation procedure requires that the signal be known to lie within some interval. The adaptive filter detailed in Sections 2.2 and 2.3 was able to accurately identify the interval in which the signal existed.

A set of relative arrival times has now been obtained. These times are now processed by the third stage to determine if they jointly indicate that their respective transients all came from the same location in space. The criterion as to whether this is, in fact, the case is given by the area of the x,y confidence ellipse. The definition of the x,y confidence ellipse is that elliptic contour within which we can be 95% certain that the transient wave front originated.

The equations that follow explicitly state how the x,y position is obtained from the relative arrival times, and how the confidence ellipse is determined. Also presented is the statistical means by which the various geophone channels are combined to obtain estimates which are more precise. The times T_i are those obtained by the cross correlation procedure just described. Note that the time T_i to the reference phone is zero.

The fundamental physical model for the original TRAPS is both simple and effective. The program is not however, in any way commited to these equations and, as will be discussed in later sections, these equations were made more sophisticated by the research efforts of this contract.

As depicted in Figure B, suppose a source originating at time T is located at position x,y and its seismic wave is received at times T_i: i = 1,...,N. Let the ranges to these sensors be denoted by r_i: i = 1,...,N. Then to first order (assuming straight line travel), we have the following equations:

$$T_i = T + \frac{1}{v}\left[(x_i-x)^2 + (y_i-y)^2\right]^{\frac{1}{2}}$$

These comprise n equations in 4 unknowns (x, y, t, v), and if N=4 they can be solved exactly. However, the arrival times in fact contain random noise perturbations which will therefore introduce error into the solution. Consequently, many observations are made and the solution which maximizes the likelihood function is chosen.

The model this becomes:

$$T_i = T + \frac{1}{v}\left[(x_i-x)^2 + (y_i-y)^2\right]^{\frac{1}{2}} + \varepsilon_i$$

where the Gaussian random process $\underset{\sim}{\varepsilon}$ has zero mean and known covariance matrix $\Lambda_{\underset{\sim}{\varepsilon}}$.

We choose x, y, T, v to minimize the form

$$Q(x,y,T,v) = \sum_{i,j}^{N} (T_i-h_i)\left[\bigwedge_{\underset{\sim}{\varepsilon}}^{-1}\right]_{ij} (T_j-h_j)$$

where

$$h_i = T + \frac{1}{v}\left[(x_i-x)^2 + (y_i-y)^2\right]^{\frac{1}{2}} \qquad i = 1,\ldots,N$$

$\bigwedge_{\underset{\sim}{\varepsilon}}^{-1}$ denotes the inverse of $\bigwedge_{\underset{\sim}{\varepsilon}}$

We now let: $\underset{\sim}{x} = (x,y,T,v)^T$

$$H_{ij}(x) = \left.\frac{\partial hi}{\partial xj}\right|_{\underset{\sim}{x}}$$

The normal equations for this problem become

$$-\sum_{ij}^{N} H_{ik}\left[\bigwedge_{\underset{\sim}{\varepsilon}}^{-1}\right]_{ij} (T_j-h_j) = 0 \qquad k = 1,\ldots,4$$

These non-linear equations are to be solved for the unknown vector $\hat{\underset{\sim}{x}}$ estimating $\underset{\sim}{x}$.

The solution is obtained as the convergent limit of the Gauss-Newton iteration scheme

$$\hat{\underset{\sim}{x}}^{(n+1)} = \hat{\underset{\sim}{x}}^{(n)} + \left[(H^T \bigwedge_{\underset{\sim}{\varepsilon}}^{-1} H)^{-1} H^T \bigwedge_{\underset{\sim}{\varepsilon}}^{-1}\right]_{\hat{\underset{\sim}{x}}^{(n)}} \underset{\sim}{r}^{(n)}$$

where $\underset{\sim}{x}^{(n)}$ denotes the value of $\underset{\sim}{x}$ at the n^{th} iteration

$$\underset{\sim}{r}^{(n)} = \underset{\sim}{T} - \underset{\sim}{h}\,(\hat{\underset{\sim}{x}}^{(n)})$$

$$\underset{\sim}{T} = (T_1, T_2, \ldots, T_N)^T$$

$\hat{\underset{\sim}{x}}^{(0)}$ is an initial guess

The iterative scheme will converge for virtually any choice of $\hat{\underset{\sim}{x}}^{(0)}$ to the same limit. In the program implemented the matrix $\bigwedge_{\underset{\sim}{\varepsilon}}$ is taken to be σ^2_{RMS} times the identity matrix.

$$\sigma^2_{RMS} = \sum_i^N \; T_i - h_i \; (\underset{\sim}{x}^{(n)}) \;^2$$

The estimator $\hat{\underset{\sim}{x}}$ has a covariance matrix given by

$$\bigwedge_{\hat{\underset{\sim}{x}}} = (H^T \bigwedge_{\underset{\sim}{\varepsilon}}^{-1} \; H)^{-1}$$

The 1σ confidence ellipse about $\hat{\underset{\sim}{x}}$ is derived from this matrix.

The above method estimates a velocity which should be taken to represent the mean horizontal group velocity from the source to the array. To a great extent the method succeeds on real data because this mean velocity is a very stable physical parameter, and its value is not assumed to be known by the program.

The equations governing the combination of observations are as follows:
Let $\hat{\underset{\sim}{x}}$ and $\hat{\underset{\sim}{y}}$ independently estimate a parameter $\underset{\sim}{\alpha}$. Let $\bigwedge_{\hat{\underset{\sim}{x}}}$ and $\bigwedge_{\hat{\underset{\sim}{y}}}$ denote their covariance matrices. Then the minimum unbiased composite linear estimator $\hat{\underset{\sim}{z}}$ is given by

$$\bigwedge_{\underset{\sim}{z}}^{-1} \cdot \hat{\underset{\sim}{z}} = \bigwedge_{\hat{\underset{\sim}{x}}}^{-1} \cdot \hat{\underset{\sim}{x}} + \bigwedge_{\hat{\underset{\sim}{y}}}^{-1} \cdot \hat{\underset{\sim}{y}}$$

$$\Lambda_{\hat{z}}^{-1} = \Lambda_{\hat{x}}^{-1} + \Lambda_{\hat{y}}^{-1}$$

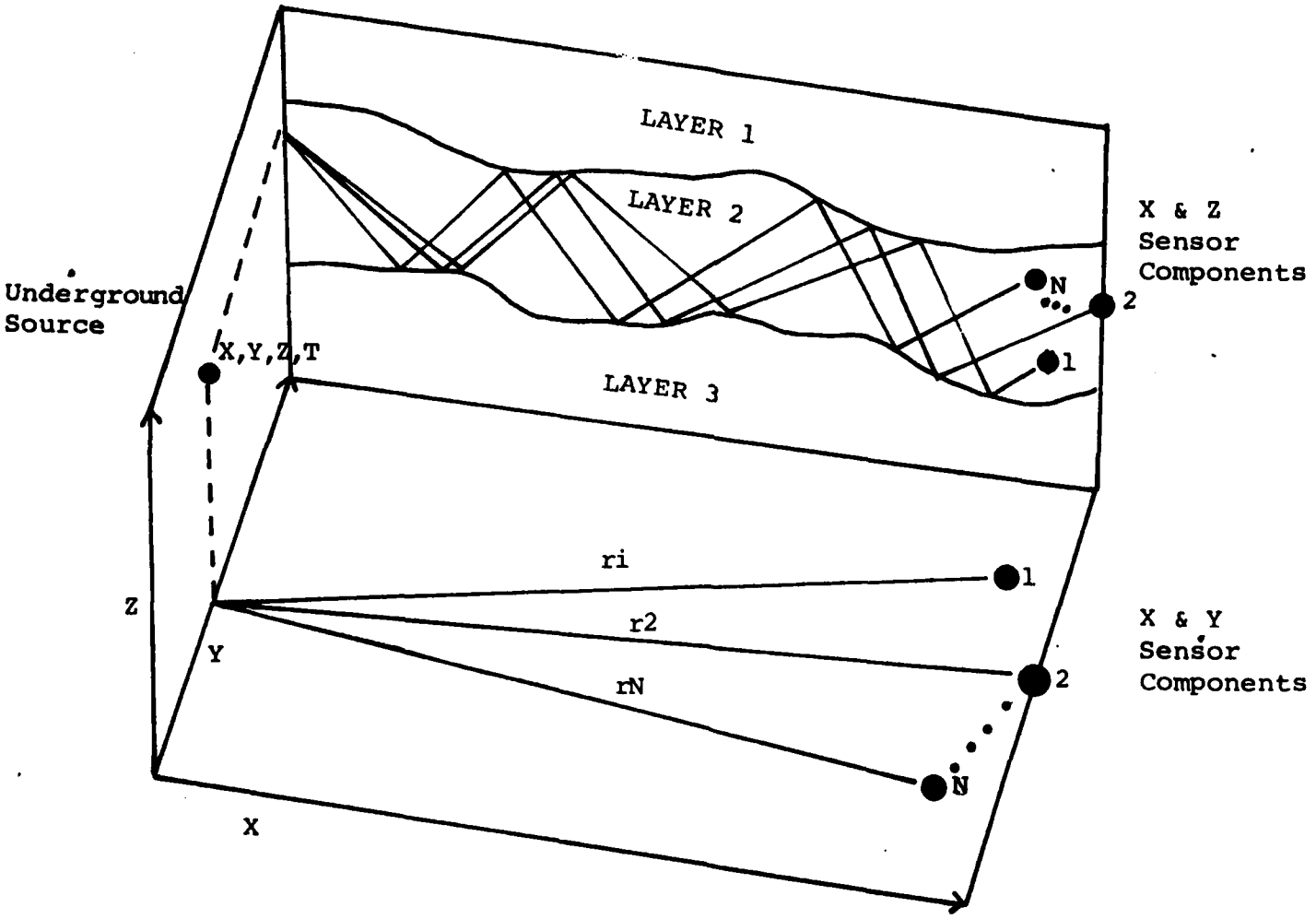


Figure A. System Geometry

APPENDIX C

APPENDIX C
INFORMATION RELATED TO
NORESS SEISMIC ARRAY DATA

1. Seismic data has been collected in one of two array configurations:

   a. A 5 element, 3 axis per element array during the time period 10 June 1983 to 5 July 1983.

   b. A 21 element, vertical axis array starting approximately 1 September 1983.

2. The data sample rate is 40 Hz and the system displacement response is depicted by Figure 1.

3. The three axis systems (Z, N, E) were located at positions 1, 2, 6, 9 and 11 used in original NORESS configuration (see Figure 2). Locations are given referenced to position #2. Position #2 is located at 60.73523° North latitude and 11.54136° East longitude. Location data are provided in Table I. In Table I and Table II, a positive X coordinate is East of the reference point and a positive Y coordinate is north of the reference point.

4. The 21 element vertical axis array was deployed at points near the points to be used in the planned DARPA/DOE/ NORSAR 1984 array. (See Figure 3). The center location (A. or Channel 1) is near the reference location (position #2) defined for the 3 axis array. Element locations, referenced to the center location are defined in Table II.

TABLE I

3 Axis Array Locations

Channel 3 3	NORESS Location	Axis	DATA TAPE CHANNEL ID NAME	X (EW)	Y (NS)	Z
1	2	Z	2Z	0	0	0
2	2	N	2N			
3	2	E	2E			
4	1	Z	1Z	-62	+271	-5
5	1	N	1N			
6	1	E	1E			
7	6	Z	6Z	-239	-348	-5
8	6	N	6N			
15	6	E	6E			
16	9	Z	9Z	-660	-407	+10
17	9	N	9N			
18	9	E	9E			
19	11	Z	BZ	-502	+1232	-25
20	11	N	BN			
21	11	E	BE			
22	(Wind Speed)		WS			

TABLE II

Vertical Axis Array Locations

Channel #	NORESS Location	Axis	DATA TAPE CHANNEL ID NAME	Location (Meters) X(EW)	Y(NS)	Z
1	Center (A_o)	Z	AØ	0	0	0
2	A_1		A1	18.2	144.6	-15.3
3	A_2		A2	124.4	-57.1	1.3
4	A_3		A3	-144.9	-66.4	-8.1
5	B_1		B1	63.5	270.5	-2.7
6	B_2		B2	316.0	-7.6	11.5
7	B_3		B3	204.9	-234.4	15.5
8	B_4		B4	-185.3	-245.2	-1
9	B_5		B5	-266.1	127.0	-12.3
10	C_1		C1	59.0	693.5	9.7
11	C_2		C2	623.8	361.0	39.7
12	C_3		C3	658.3	-207.9	56.5
13	C_4		C4	202.1	-671.6	6.8
14	C_5		C5	-360.6	-624.6	-4.1
15	C_6		C6	-770.5	-82.9	10.0
16	C_7		C7	-454.9	571.2	-26.4
17	D_1		D1	138.4	1476.4	5.9
18	D_2		D2	1060.2	1100.5	67.6
19	D_4		D4	1225.1	-933.4	81.6
20	D_5		D5	-412.2	-1466.6	45.3
21	D_7	Z	P7	-1347.2	-734.7	32.3
22	(Wind Speed)		WS			

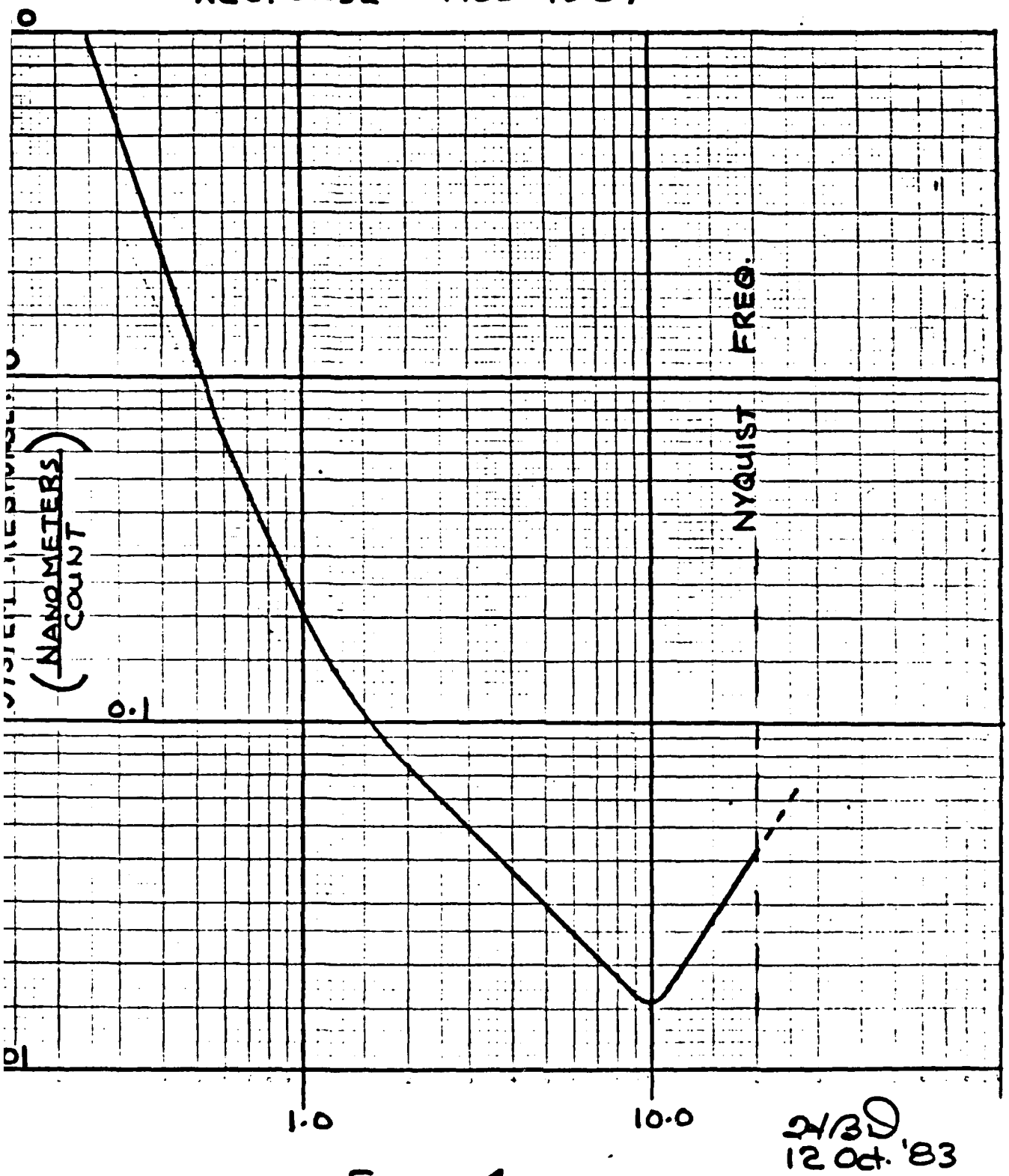

NORESS SP ARRAY
RESPONSE 1983-1984
0
0.1
(NANOMETERS)
COUNT
NYQUIST FREQ.
1.0
10.0
2/3D
12 Oct. '83

FIGURE 1

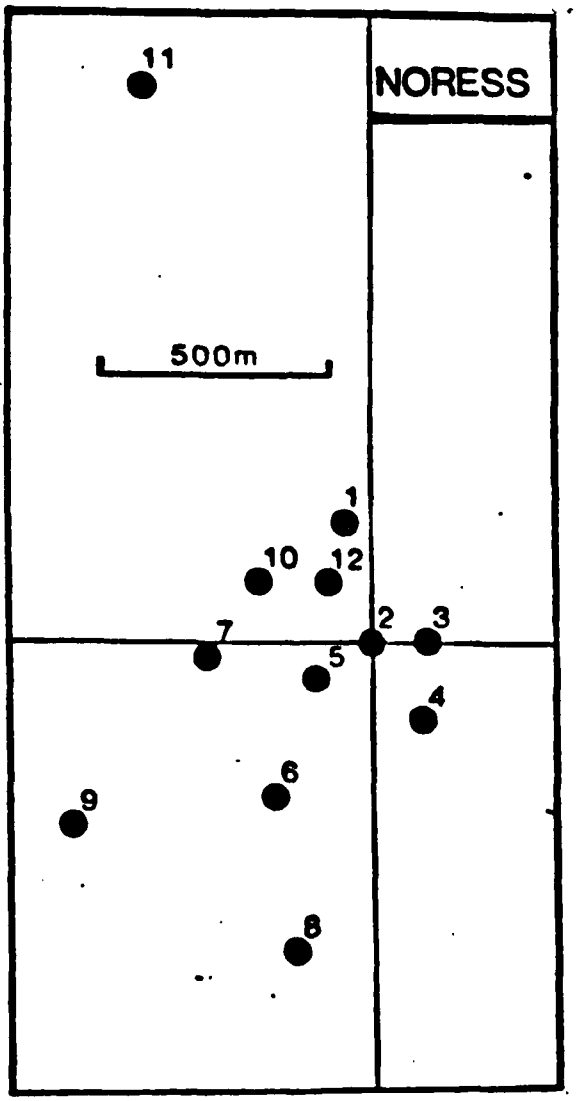


FIGURE 2

TABLES

Table 1. Periodogram Resolution as a Function of SNR

SNR	MEAN OF R_N	STANDARD DEVIATION OF R_N
∞	0.031374	±0.0
20 dB	0.031374	±0.000082
10 dB	0.031359	±0.000258
0 dB	0.031328	±0.000798
-10 dB	0.030989	±0.002362

Table 2. NORSAR/NORESS Data Description
(see Appendix C for further detail)

Processing Event No.	Date	Time-Arrival	Geogr.	Coord.	Event Description
1	June 12	02.44.05.7	49N	78E	EASTERN KAZAKH SSR
2	June 12	12.07.05.2	41N	30E	TURKEY
3	June 10	02.20.36.9	76N	125E	LAPTEV SEA
4	June 13	15.13.30.0	61N	15E	SWEDEN
5	June 24	03.03.31.0	51N	78E	EASTERN KAZAKH SSR
6	June 29	20.59.32.6	60N	13E	SWEDEN
7	June 29	15.08.36.9	67N	20E	SWEDEN
8	June 30	06.28.23.3	61N	8E	SOUTHERN NORWAY
1' (day 272)	Sept 29	05.03.22.72	63.9N	17.5E	SWEDEN - EARTHQUAKE
2' (day 276)	Oct 3	09.50.40.68	59.7N	13.25	SWEDEN - EARTHQUAKE
3'	day 284	14.30.36.00	59.7N	10.7E	MINING EXPLOSION
4'	day 294	12.11.00.00	59.3N	6.9E	MINING EXPLOSION
5'	day 305	10.35.00.00	60.5N	5.2E	NAVY UNDERWATER EXPLOSION
6'	day 305	12.20.00.00	60.4N	5.3E	BLAST IN WATER
7'	day 305	12.43.00.00	60.0N	4.5E	NAVY UNDERWATER EXPLOSION
8'	day 312	00.49.32.00	50.7N	5.4E	BELGIUM - EARTHQUAKE

Table 3. Original TRAPS Bearing Results

Event No.	1	2'	3'	4'	5'	6'	7'	8'
Azimuth Accuracy	±1°	±.05°	±1°	±1°	±1°	±3°	±3°	±.05°

Table 4. TRAPS Bearing Results with CAARMA Model

Event No.	1	2'	3'	4'	5'	6'	7'	8'
Azimuth Accuracy	±1°	±.05°	±1°	±1°	±1°	±1°	±1°	±.05°

Table 5. TRAPS Bearing Results with CAARMA Model and Improved Time Delay Method

Event No.	1	2'	3'	4'	5'	6'	7'	8'
Azimuth Accuracy	±.05°	±.05°	±1°	±1°	±1°	±1°	±1°	±.05°

Table 6. TRAPS Bearing Results with CAARMA Model, Improved Time Delay & ML Feedback Scheme

Event No.	1	2'	3'	4'	5'	6'	7'	8'
Azimuth Accuracy	±.01°	±.01°	±.02°	±.02°	±.01°	±.05°	±.05°	±.01°